De-Shuang Huang Donald C. Wunsch II
Daniel S. Levine Kang-Hyun Jo (Eds.)

Advanced Intelligent Computing Theories and Applications

With Aspects of Theoretical and Methodological Issues

LNCS 5226

4th International Conference on Intelligent Computing, ICIC 2008
Shanghai, China, September 2008
Proceedings

Springer

Lecture Notes in Computer Science 4226

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison
Lancaster University, UK

Takeo Kanade
Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler
University of Surrey, Guildford, UK

Jon M. Kleinberg
Cornell University, Ithaca, NY, USA

Friedemann Mattern
ETH Zurich, Switzerland

John C. Mitchell
Stanford University, CA, USA

Moni Naor
Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz
University of Bern, Switzerland

C. Pandu Rangan
Indian Institute of Technology, Madras, India

Bernhard Steffen
University of Dortmund, Germany

Madhu Sudan
Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos
University of California, Los Angeles, CA, USA

Doug Tygar
University of California, Berkeley, CA, USA

Moshe Y. Vardi
Rice University, Houston, TX, USA

Gerhard Weikum
Max-Planck Institute of Computer Science, Saarbruecken, Germany

Roland T. Mittermeir (Ed.)

Informatics Education – The Bridge between Using and Understanding Computers

International Conference in Informatics in Secondary Schools – Evolution and Perspectives, ISSEP 2006
Vilnius, Lithuania, November 7-11, 2006
Proceedings

Springer

Volume Editor

Roland T. Mittermeir
Institut für Informatik-Systeme
Universität Klagenfurt
9020 Klagenfurt, Austria
E-mail: roland@isys.uni-klu.ac.at

Library of Congress Control Number: 2006935052

CR Subject Classification (1998): K.3, K.4, J.1, K.8

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

ISSN 0302-9743
ISBN-10 3-540-48218-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-48218-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2006
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11915355 06/3142 5 4 3 2 1 0

Preface

Although the school system is subject to specific national regulations, didactical issues warrant discussion on an international level. This applies specifically to informatics didactics. In contrast to most other scientific disciplines, informatics undergoes substantial technical and scientific changes and shifts of paradigms even at the basic level taught in secondary school. Moreover, informatics education is under more stringent observation from parents, potential employers, and policy makers than other disciplines. It is considered to be a modern discipline. Hence, being well-educated in informatics seemingly ensures good job perspectives. Further, policy makers pay attention to informatics education, hoping that a young population well-educated in this modern technology will contribute to the future wealth of the nation. But are such high aspirations justified? What should school aim at in order to live up to such expectations?

ISSEP 2005, the 1st International Conference on Informatics in Secondary Schools – Evolution and Perspectives already showed that informatics teachers have to bridge a wide gap [1, 2]. On one hand, they have to show the inherent properties that informatics (or computer science) can contribute to general education. On the other hand, they are to make pupils computer literate. Under the constraint of limited time available for instruction, these different educational aims come into conflict.

Computer-supported teaching or eLearning is to be considered distinct from informatics education. However, in many countries, informatics teachers still have to support the eTeaching activities of their colleagues. They might even be the only ones to support eLearning. But even in situations where teachers of other subject areas are sufficiently computer literate to use computer support in their own courses, they will expect students to arrive already technically prepared by informatics courses.

Considering this spectrum, the program of the 2nd International Conference on Informatics in Secondary Schools – Evolution and Perspectives, ISSEP 2006, was mainly structured into discussions on what and how to teach. Those aiming at educating "informatics proper" by showing the beauty of the discipline, hoping to create interest in a later professional career in computing, will give answers different from the opinion of those who want to familiarize pupils with the basics of ICT in order to achieve computer literacy for the young generation. Addressing eLearning aspects as seen from the perspective of informatics didactics are another only moderately related set of issues. This spread of topics raises the question of what is a proper examination to assess students' performance. Furthermore, one has to see that school-informatics is still (and will remain in the foreseeable future) a subject in transition. Hence, teacher's education was also in the focus of ISSEP 2006. Consequently, the selection of papers contained in these proceedings address the topics just mentioned. Further discussions of these and related topics are covered in "*Information Technologies at Schools*" [3], the remaining part of the proceedings of ISSEP 2006.

The 29 papers contained in this volume were selected out of a total of 204 submissions and invited contributions. The accompanying volume [3] contains 70

scientific papers. Some 50 rather school-practical contributions targeted for the "Lithuanian Teachers Session" are made available on a CD (in Lithuanian) [4]. Each scientific paper was reviewed by at least three members of the Program Committee. The reviewing process and the ensuing discussion were fully electronic.

This volume, although consisting mainly of contributed papers, is nevertheless the result of an arrangement of papers aiming in their final versions to contribute to the specific facet of the program they were accepted for. The remainder of this preface shows how they contribute to the various facets of the conference.

The core of papers contained in this volume center on the *tension between* making pupils familiar with the *fundamental ideas* upon which the discipline of informatics rests, following an aim similar to education in physics or chemistry, *and ICT* or *computer literacy* instruction. *Dagienė, Dzemyda, and Sapagovas* open this series of papers by reporting the development of informatics education in Lithuania. Due to the political and related social changes in this country, the differences as well as the similarities to developments in other countries are of particular interest. The following papers address the issue of familiarizing students with informatics fundamentals from very different angles. *Kalaš* describes a course where a Logo-platform supports explorative learning. Specific focus is given on (behavioral) modeling, visualizations of fractions, and biological growth. From the different examples, students can identify structure and finally develop algorithmic problem-solving skills. *Hromkovič* describes his approach of relating the beauty of informatics to students attending a course supplementary to general school education. The paper presents the rationale behind kindling pupils' interest in informatics as a distinct science and explains related didactical aspects. Still at the "high end" of informatics education is the extra-curricular program described by *Yehezkel and Haberman*. Departing from the assumption that in general teachers lack experience and credibility as professional software developers, the authors developed a program where graduates from secondary level schools work on a real project under the mentorship of professional software developers.

In order not to lose focus, the paper by *Szlávi and Zsakó* contrasts two aspects of informatics education: the aim to teach future users of IT-systems and the aim to educate future programmers. The presentation is stratified according to educational aims attainable at particular age levels. In spite of the contrasts highlighted by this paper, *Antonitsch* shows that there are bridges between teaching applications and teaching fundamental concepts. His paper, based on a database application, can be seen as a continuation of bridging approaches reported by Voss (departing from text-processing) and by Antonitsch (departing from spreadsheet-modeling) at ISSEP 2005 [1]. Raising the student's curiosity by showing informatics' concepts in such varied disciplines as mathematics, biology, and art is the subject of *Sendova's* paper. Her approach ensures a low entrance-barrier, but still leads to elementary algorithmic and programming skills.

Clark and Boyle analyze the developments in English schools. Although the British school system differs quite a bit from its continental counterpart, the trends identified by analyzing developments form 1969 onwards find their analogs in most other countries that introduced formal informatics education. Special consideration might be given to their projection into the future. Currently, we still live in a situation where most parents are not computer literate. But this deficiency will gradually vanish

during the years to come. How should school react to a situation when pupils become computer literate following their parents' or their peers' IT-related activities?

The selection of papers on fundamentals is terminated by the work of *Haberman*. It directly leads into both the section on programming and the section on ICT. Specifically, Habermann focuses on the educational milieu and on a gap in perception as to what computing (informatics) is all about. Perhaps resolving this terminological issue, as it has been resolved in distinguishing between learning basic arithmetic (calculating) and mathematics, might solve some public misunderstandings and related problems.

The papers in the initial part of the proceedings focus on the question of *"What to teach?"* To a varying extent they address this question in the context of constrained time to provide the respective qualifications to students. The succeeding set of papers addresses didactical issues of a core aspect of instruction about informatics proper, i.e., *programming and algorithms*. The key question there is: *"How to teach (programming)?*" This part of the proceedings is opened by *Hubwieser*, who explains how object-oriented programming was introduced in the context of a situation where the overall time for informatics education was restricted with respect to initial plans. While Hubwieser's approach for Bavaria foresees a focus on object-oriented software, the paper of *Weigend* addresses three basic issues related to the problem that the capability of performing a task (procedural intuition) is still insufficient for being able to formulate the individual steps necessary to conduct this task (e.g., to write a program). A Python-based system is proposed to overcome this mental barrier. But the problem of finding an appropriate algorithm has many facets. *Ginat* shows the dangers of focusing exclusively on the mainstream strategy of divide-and-conquer for solving algorithmic problems. He points to examples where a global perspective is necessary for obtaining a correct and efficient solution. One might perceive of this paper as a counterpoint to mainstream teaching. It makes teachers and students aware that problem solving needs a rich repertoire of strategies and recipes. There is no once-and-for-all solution.

Kurebayashi, Kamada, and Kanemune report on an experiment involving 14- to 15-year-old pupils in programming simple robots. The authors' approach combines playful elements with serious programming. It is interesting to see that their experiments showed the particular usefulness of this approach for pupils with learning deficiencies.

The master class in software engineering described by *Verhoeff* attaches well to the approaches followed by Hromkovič and by Yehezkel and Haberman. Pupils are invited to this extra-curricular master course which is co-operatively taught at school and at university. The approach of having students complete a small programming project in a professional manner is described in detail. Another concept of a pre-university course to foster algorithmic thinking is described by *Futschek*. He gives three specific examples that can be studied with young people transiting from school to university.

Laucius presents a socio-linguistic issue. While English is the language of computing, one cannot assume too much previous knowledge of this language with pupils if – as for most countries – English is a foreign language. In the case that the local language uses even a different character set, problems are aggravated. Hence, this issue is addressed in several papers by Lithuanian authors. The critical question,

however, might be how far one should go in localizing computer science. The "foreign" language is definitely a hurdle. However, controlled use of the foreign language allows one to clearly separate between object-language and meta-language. To close this circle, the paper by *Salanci* returns to object-oriented programming by presenting an approach for a very smooth, stepwise introduction to working with software-objects.

Papers on *ICT instruction* constitute the ensuing part of the proceedings. They can be seen as a companion to the discussion presented so far. *Micheuz* discusses the selection of topics to be covered in ICT lessons from the perspective of an increasing autonomy within a school system that is at the same time burdened by new constraints (reductions) on the number of courses it may offer. It is interesting to note that an "invisible hand" managed to ensure convergence of the topics finally covered. SeungWook Yoo et al. explain how adoption of model curricula helped to solve problems in informatics education in Korea. *Syslo and Kwiatkowska* conclude this set of papers by noting that the link between mathematics education and informatics education is essentially bi-directional. However, in most current school-books only one of these directions is made explicit. The paper presents some examples where mathematics education could benefit from adopting concepts of informatics.

The widely discussed topics of school informatics addressed so far need context. This context is to be found in the relationships between (maturity) *exams* and informatics instruction, as addressed by *Blonskis and Dagienė*. With the wealth of extra-curricular activities and competitions such as the International Olympiad in Informatics, the question of proper scoring, notably the issue of arriving at a scoring scheme that is not de-motivating to those who are not victorious, becomes of interest. *Kemkes, Vasiga, and Cormack* propose a weighting scheme for automatic test assessments. Their results are generally applicable in situations where many programs are to be graded in a standardized manner and assessments are strictly functionality-based.

Teachers' education and school development is a different contextual aspect. *Markauskaite, Goodwin, Reid, and Reimann* address the challenges of providing good ICT courses for pre-service teachers. The phenomenon of different pre-knowledge is a well-known didactical problem when familiarizing pupils with ICT concepts. This problem is aggravated in educating future teachers. Some of them will be recent graduates – possibly even with moderate motivation to learn (and use) ICT – while others might look back on a non-educational professional career that may have involved already substantial contact with computing. Special recommendations of how to cope with this problem are given. The focus of *Butler, Strohecker, and Martin* is, in contrast, on teachers that are already experienced in their profession but follow a rather traditional style of teaching. By entering a collaborative project with their pupils, constructivist teaching principles can be brought into practice. Moreover, changes in the teacher's and students' roles become noticeable. The ensuing open style of learning is appreciated by all parties of the school system and the approach spreads quite well throughout Ireland.

The proceedings conclude with contributions related to *eLearning. Kahn, Noss, Hoyles, and Jones* report on their environment supporting layered learning. This environment allows pupils to construct games where the outcome depends on proper application of physical principles by the student-players. Enriching the model, one

can increase the depth concerning physics instruction. But the layered approach also allows one to manipulate games in such a way that finally (fragments of) programs can be written by the students.

ePortfolios currently attract a lot of attention in didactical circles. *Hartnell-Young's* paper is a worthwhile contribution to this debate, as it presents results from four schools and a special cluster, each with different aims targeted specifically for the student population to be supported. In any case, scope and aspirations were limited but results were encouraging. The paper might well serve as a warning for those who believe a particular ePortfolio can satisfy all those goodies portfolios can support in principle. Remaining at the level of meta-cognition, *Giuseppe Chiazzese et a*l. present a tool that makes students aware of the activities (partly subconsciously) perfomed while surfing the Web. Pursuing these ideas further, a transition from computer literacy to Web literacy might be finally achieved at school.

The proceedings conclude with two papers referring to aspects of internationalizing and localizing instructional software. *Targamadzė and Cibulskis* describe the development of the Lithuanian Distance Education Network, a project pursued on the European international level. *Jevsikova* provides a detailed list of issues to be observed when one prepares courseware intended for use on an international level.

A conference like this is not possible without many hands and brains working for it and without the financial support of graceful donors. Hence, I would like to thank particular in the General Chair and the members of the Program Committee, notably those who were keen to review late arrivals as well as those colleagues who provided additional reviews. Special thanks are due to the Organizing Committee led by Roma Žakaitienė and Gintautas Dzemyda. Karin Hodnigg deserves credit for operating the electronic support of the submission and reviewing process, Annette Lippitsch for editorial support for these proceedings.

The conference was made possible due to the support of several sponsors whose help is gratefully acknowledged. Printing and wide distribution of its two volumes of proceedings were made possible due to a substantial contribution by the Goverment of the Republic of Lithuania, and the Ministry of Education and Science of Lithuania. Finally, hosting of the conference by Seimas, the Lithuanian Parliament, is gratefully acknowledged.

November 2006 Roland Mittermeir

1. Mittermeir R.: From Computer Literacy to Informatics Fundamentals, Proc. ISSEP 2005 (part 1), LNCS 3422, Springer Verlag, Berlin, Heidelberg, 2005.
2. Micheuz P., Antonitsch P., Mittermeir R.: Informatics in Secondary Schools – Evolution and Perspectives: Innovative Concepts for Teaching Informatics, Proc ISSEP 2005 (part 2), Ueberreuter Verlag, Wien, March 2005.
3. Dagienė V., Mittermeir R.: Information Technologies at School; Publ: TEV, Vilnius, October, 2006
4 Dagienė V., Jasutienė E., Rimkus M.: Informacinės technologijos mokykloje. (in Lithuanian), CD, available from the conference webpage http://ims.mii.lt/imrp.

Organization

ISSEP 2006 was organized by the Institute of Mathematics and Informatics, Lithuania.

ISSEP 2006 Program Committee

Valentina Dagienė (Chair)	Institute of Mathematics and Informatics, Lithuania
Roland Mittermeir (Co-chair)	Universität Klagenfurt, Austria
Andor Abonyi-Tóth	Eötvös Loránd University, Hungary
Iman Al-Mallah	Arab Academy for Science & Technology and Maritime Transport, Egypt
Juris Borzovs	University of Latvia, Latvia
Laszlo Böszörmenyi	Universität Klagenfurt, Austria
Roger Boyle	University of Leeds, UK
Norbert Breier	Universität Hamburg, Germany
Giorgio Casadei	University of Bologna, Italy
David Cavallo	Massachusetts Institute of Technology, USA
Mike Chiles	Western Cape Education Department, South Africa
Martyn Clark	University of Leeds, UK
Bernard Cornu	CNED-EIFAD (Open and Distance Learning Institute), France
Zide Du	China Computer Federation, China
Steffen Friedrich	Technische Universität Dresden, Germany
Karl Fuchs	Universität Salzburg, Austria
Patrick Fullick	University of Southampton, UK
Gerald Futschek	Technische Universität Wien, Austria
David Ginat	Tel-Aviv University, Israel
Juraj Hromkovič	Swiss Federal Institute of Technology Zürich, Switzerland
Peter Hubwieser	Technische Universität München, Germany
Feliksas Ivanauskas	Vilnius University, Lithuania
Ivan Kalaš	Comenius University, Slovakia
Susumu Kanemune	Hitotsubashi University, Japan
Ala Kravtsova	Moscow Pedagogical State University, Russia
Nancy Law	The University of Hong Kong, Hong Kong
Lauri Malmi	Helsinki University of Technology, Finland
Krassimir Manev	Sofia University, Bulgaria
Peter Micheuz	Universität Klagenfurt and Gymnasium Völkermarkt, Austria

Paul Nicholson	Deakin University, Australia
Nguyen Xuan My	Hanoi National University, Vietnam
Richard Noss	London Knowledge Lab, Institute of Education,University of London, UK
Sindre Røsvik	Giske Municipality, Norway
Ana Isabel Sacristan	Center for Research and Advanced Studies (Cinvestav), Mexico
Tapio Salakoski	Turku University, Finland
Sigrid Schubert	Universität Siegen, Germany
Aleksej Semionov	Moscow Institute of Open Education, Russia
Carol Sperry	Millersville University, USA
Oleg Spirin	Zhytomyr Ivan Franko University, Ukraine
Maciej M.Syslo	University of Wrocław, Poland
Aleksandras Targamadzė	Kaunas University of Technology, Lithuania
Laimutis Telksnys	Institute of Mathematics and Informatics, Lithuania
Armando Jose Valente	State University of Campinas, Brazil
Tom Verhoeff	Eindhoven University of Technology, Netherlands
Aleksander Vesel	University of Maribor, Slovenia
Anne Villems	University of Tartu, Estonia

Additional Reviewers

Peter Antonitsch	Universität Klagenfurt and HTL Mössingerstr., Klagenfurt, Austria
Mats Daniels	Uppsala University, Sweden
Karin Hodnigg	Universität Klagenfurt, Austria
Kees Huizing	Eindhoven University of Technology, Netherlands
Toshiyuki Kamada	Aichi University of Education, Japan
Shuji Kurebayashi	Shizuoka University, Japan
Ville Leppänen	University of Turku, Finland
Don Piele	University of Wisconsin, USA

Organizing Committee

Roma Žakaitienė (Chair)	Ministry of Education and Science of the Republic of Lithuania
Gintautas Dzemyda (Co-chair)	Institute of Mathematics and Informatics, Lithuania
Vainas Brazdeikis	Centre of Information Technologies of Education, Lithuania

Ramūnas Čepaitis	Chancellery of the Seimas of the Republic of Lithuania
Karin Hodnigg	Universität Klagenfurt, Austria
Annette Lippitsch	Universität Klagenfurt, Austria
Jonas Milerius	Chancellery of the Seimas of the Republic of Lithuania
Algirdas Monkevicius	Seimas of the Republic of Lithuania
Gediminas Pulokas	Institute of Mathematics and Informatics, Lithuania
Alfonsas Ramonas	Chancellery of the Seimas of the Republic of Lithuania
Modestas Rimkus	Institute of Mathematics and Informatics, Lithuania
Danguolė Rutkauskienė	Kaunas Technology University, Lithuania
Elmundas Žalys	Publishing House TEV, Lithuania
Edmundas Žvirblis	Information Society Development Committee under the Government of the Republic of Lithuania

Main Sponsors

ISSEP 2006 and the publication of its proceedings were supported by the Government of the Republic of Lithuania, and Ministry of Education and Science of the Republic of Lithuania.

Table of Contents

The Spectrum of Informatics Education

Evolution of the Cultural-Based Paradigm for Informatics Education in Secondary Schools – Two Decades of Lithuanian Experience 1

Discovering Informatics Fundamentals Through Interactive Interfaces for Learning 13

Contributing to General Education by Teaching Informatics 25

Bridging the Gap Between School Computing and the "Real World" 38

Programming Versus Application 48

Databases as a Tool of General Education 59

Handling the Diversity of Learners' Interests by Putting Informatics Content in Various Contexts 71

Computer Science in English High Schools: We Lost the S, Now the C Is Going 83

Teaching Computing in Secondary Schools in a Dynamic World: Challenges and Directions 94

Teaching Algorithmics and Programming

Functions, Objects and States: Teaching Informatics in Secondary Schools 104

From Intuition to Programme 117

On Novices' Local Views of Algorithmic Characteristics 127

Learning Computer Programming with Autonomous Robots 138

A Master Class Software Engineering for Secondary Education 150

Algorithmic Thinking: The Key for Understanding Computer Science 159

Issues of Selecting a Programming Environment for a Programming Curriculum in General Education 169

Object-Oriented Programming at Upper Secondary School for Advanced Students 179

The Role of ICT-Education

Informatics Education at Austria's Lower Secondary Schools Between Autonomy and Standards 189

Development of an Integrated Informatics Curriculum for K-12 in Korea 199

Contribution of Informatics Education to Mathematics Education in Schools 209

Exams and Competitions

Evolution of Informatics Maturity Exams and Challenge for Learning Programming 220

Objective Scoring for Computing Competition Tasks 230

Teacher Education and School Development

Modelling and Evaluating ICT Courses for Pre-service Teachers: What Works and How It Works? 242

Sustaining Local Identity, Control and Ownership While Integrating Technology into School Learning 255

eLearning

Designing Digital Technologies for Layered Learning 267

ePortfolios in Australian Schools: Supporting Learners' Self-esteem, Multiliteracies and Reflection on Learning 279

Metacognition in Web-Based Learning Activities 290

Development of Modern e-Learning Services for Lithuanian Distance Education Network LieDM 299

Localization and Internationalization of Web-Based Learning Environment 310

Author Index 319

Some Observations on Two-Way Finite Automata with Quantum and Classical States*

Daowen Qiu

Department of Computer Science, Zhongshan University, Guangzhou 510275, China
issqdw@mail.sysu.edu.cn

Abstract. *Two-way finite automata with quantum and classical states* (2qcfa's) were introduced by Ambainis and Watrous. Though this computing model is more restricted than the usual *two-way quantum finite automata* (2qfa's) first proposed by Kondacs and Watrous, it is still more powerful than the classical counterpart. In this note, we focus on dealing with the operation properties of 2qcfa's. We prove that the Boolean operations (intersection, union, and complement) and the reversal operation of the class of languages recognized by 2qcfa's with error probabilities are closed; also, we verify that the catenation operation of such class of languages is closed under certain restricted condition. The numbers of states of these 2qcfa's for the above operations are presented. Some examples are included, and $\{xx^R | x \in \{a,b\}^*, \#_x(a) = \#_x(b)\}$ is shown to be recognized by 2qcfa with one-sided error probability, where x^R is the reversal of x, and $\#_x(a)$ denotes the a's number in string x.

1 Introduction

Quantum computers—the physical devices complying with quantum mechanics were first suggested by Benioff [3] and Feynman [9] and then formalized further by Deutsch [6]. A main goal for exploring this kind of model of computation is to clarify whether computing models built on quantum physics can surpass classical ones in essence. Actually, in 1990's Shor's quantum algorithm for factoring integers in polynomial time [19] and afterwards Grover's algorithm of searching in database of size n with only $O(\sqrt{n})$ accesses [11] have successfully shown the great power of quantum computers. Since then great attention has been payed to this intriguing field [12], and, clarifying the power of some fundamental models of quantum computation is of interest [12].

(qfa's) can be thought of theoretical models of quantum computers with finite memory [12]. With the rise of exploring quantum computers, this kind of theoretical models was firstly studied by Moore and Crutchfield [17], Kondacs and Watrous [16], and then Ambainis and Freivalds

* This work was supported by the National Natural Science Foundation under Grant 90303024 and Grant 60573006, the Research Foundation for the Doctorial Program of Higher School of Ministry of Education under Grant 20050558015, and Program for New Century Excellent Talents in University (NCET) of China.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1–8, 2008.
© Springer-Verlag Berlin Heidelberg 2008

[1], Brodsky and Pippenger [5], and the other authors (e.g., see [12]). The study of qfa's is mainly divided into two ways: one is [illegible] (1qfa's) whose tape heads move one cell only to right at each evolution, and the other [illegible] (2qfa's), in which the tape heads are allowed to move towards right or left, or to be stationary. In terms of the measurement times in a computation, 1qfa's have two types: [illegible] 1qfa's (MO-1qfa's) initiated by Moore and Crutchfield [17] and [illegible] 1qfa's (MM-1qfa's) studied firstly by Kondacs and Watrous [16].

MO-1qfa's mean that at every computation there is only a measurement at the end of computation, whereas MM-1qfa's represent that measurement is performed at each evolution. The class of languages recognized by MM-1qfa's with bounded error probabilities strictly bigger than that by MO-1qfa's, but both MO-1qfa's and MM-1qfa's recognize proper subclass of regular languages with bounded error probabilities (e.g., see [16,1,5,17]). On the other hand, the class of languages recognized by MM-1qfa's with bounded error probabilities is not closed under the binary Boolean operations [5,4], and by contrast MO-1qfa's satisfy the closure properties of the languages recognized with bounded error probabilities under binary Boolean operations [5,4].

A more powerful model of quantum computation than its classical counterpart is 2qfa's that were first studied by Kondacs and Watrous [16]. As is well known, classical two-way finite automata have the same power as one-way finite automata for recognizing languages [14]. Freivalds [10] proved that [illegible] (2pfa's) can recognize non-regular language $L_{eq} = \{a^n b^n | n \in \mathbf{N}\}$ with arbitrarily small error, but it was verified to require exponential expected time [13]. (In this paper, $\mathbf{N}$ denotes the set of natural numbers.) Furthermore, it was demonstrated that any 2pfa's recognizing non-regular languages with bounded error probabilities need take exponential expected time [7,15]. In 2qfa's, a sharp contrast has arisen, as Kondacs and Watrous [16] proved that L_{eq} can be recognized by some 2qfa's with one-sided error probability in linear time.

In 2002, Ambainis and Watrous [2] proposed a different two-way quantum computing model—[illegible] (2qcfa's). In this model, there are both quantum states and classical states, and correspondingly two transfer functions: one specifies unitary operator or measurement for the evolution of quantum states and the other describes the evolution of classical part of the machine, including the classical internal states and the tape head. This device may be simpler to implement than ordinary 2qfa's, since the moves of tape heads of 2qcfa's are classical. In spite of the existing restriction, 2qcfa's have more power than 2pfa's. Indeed, 2qcfa's can recognize all regular languages with certainty, and particularly, they [2] proved that this model can also recognize non-regular languages $L_{eq} = \{a^n b^n | n \geq 1\}$ and palindromes $L_{pal} = \{x \in \{a, b\}^* | x = x^R\}$, where notably the complexity for recognizing L_{eq} is polynomial time in one-sided error. However, no 2pfa can recognize L_{pal} with bounded error in any amount of time [8]. Therefore, this is an interesting and more practicable model of quantum computation.

Operations of finite automata are of importance [14] and also interest in the framework of quantum computing [5,17,4]. Our goal in this note is to deal with the operation properties of 2qcfa's. We investigate some closure properties of the class of languages recognized by 2qcfa's, and we focus on the binary Boolean operations, reversal operation, and catenation operation. Notwithstanding, we do not know whether or not these properties hold for the ordinary 2qfa's without any restricted condition, and we would like to propose them as an open problem (As the author is aware, the main problem to be overcome is how to preserve the unitarity of the constructed 2qfa's without any restricted condition).

The remainder of the paper is organized as follows. In Section 2 we introduce the definition of 2qcfa's [2] and further present some non-regular languages recognized by 2qcfa's. Section 3 deals with operation properties of 2qcfa's, including intersection, union, complement, reversal, and catenation operations; also, some examples are included as an application of these results derived. Finally, some remarks are made in Section 4.

Due to the limited space, the proofs in this note are left out and the details are referred to [18].

2 Definition of 2qcfa's and Some Remarks

In this section, we recall the definition of 2qcfa's and some related results [2], from which we derive some further outcomes.

A 2qcfa M consists of a 9-tuple $M = (Q, S, \Sigma, \Theta, \delta, q_0, s_0, S_{acc}, S_{rej})$ where Q and S are finite state sets, representing quantum states and classical states, respectively, Σ is a finite alphabet of input, $q_0 \in Q$ and $s_0 \in S$ denote respectively the initial quantum state and classical state, $S_{acc}, S_{rej} \subseteq S$ represent the sets of accepting and rejecting, respectively, Θ and δ are the functions specifying the behavior of M regarding quantum portion and classical portion of the internal states, respectively.

For describing Θ and δ, we further introduce related notions. We denote $\Gamma = \Sigma \cup \{¢, \$\}$, where ¢ and \$ are respectively the left end-marker and right end-marker. $l_2(Q)$ represents the Hilbert space with the corresponding base identified with set Q. Let $\mathcal{U}(l_2(Q))$ and $\mathcal{M}(l_2(Q))$ denote the sets of unitary operators and orthogonal measurements over $l_2(Q)$, respectively. An orthogonal measurement over $l_2(Q)$ is described by a finite set $\{P_j\}$ of projection operators on $l_2(Q)$ such that $\sum_j P_j = I$ and $P_i P_j = \begin{cases} P_j, i = j, \\ O, \ i \neq j, \end{cases}$ where I and O are identity operator and zero operator on $l_2(Q)$, respectively. If a superposition state $|\psi\rangle$ is measured by an orthogonal measurement described by set $\{P_j\}$, then

1. the result of the measurement is j with probability $\|P_j|\psi\rangle\|^2$ for each j,
2. and the superposition of the system collapses to $P_j|\psi\rangle / \|P_j|\psi\rangle\|$ in case j is the result of measurement.

Θ and δ are specified as follows. Θ is a mapping from $S\backslash(S_{acc} \cup S_{rej}) \times \Gamma$ to $\mathcal{U}(l_2(Q)) \cup \mathcal{M}(l_2(Q))$, and δ is a mapping from $S\backslash(S_{acc} \cup S_{rej}) \times \Gamma$ to $S \times \{-1, 0, 1\}$. To be more precise, for any pair $(s, \sigma) \in S\backslash(S_{acc} \cup S_{rej}) \times \Gamma$,

1. if $\Theta(s,\sigma)$ is a unitary operator U, then U performs the current superposition of quantum states and then it evolves into new superposition, while $\delta(s,\sigma) = (s^{'},d) \in S\times\{-1,0,1\}$ makes the current classical state s become $s^{'}$, together with the tape head moving in terms of d (moving right one cell if $d=1$, left if $d=-1$, and being stationary if $d=0$), for which in case $s^{'} \in S_{acc}$, the input is accepted, and in case $s^{'} \in Q_{rej}$, the input rejected;
2. if $\Theta(s,\sigma)$ is an orthogonal measurement, then the current quantum state, say $|\psi\rangle$, is naturally changed to quantum state $P_j|\psi\rangle/\|P_j|\psi\rangle\|$ with probability $\|P_j|\psi\rangle\|^2$ in terms of the measurement, and in this case, $\delta(s,\sigma)$ is instead a mapping from the set of all possible results of the measurement to $S\times\{-1,0,1\}$. The further details are referred to [2,18].

A language L over alphabet Σ^* is called to be recognized by 2qcfa M with bounded error probability ϵ if $\epsilon \in [0,1/2)$, and for any $x\in L$, $P_{acc}^{(M)}(x) \geq 1-\epsilon$, for any $x \in L^c = \Sigma^*\backslash L$, $P_{rej}^{(M)}(x) \geq 1-\epsilon$, where $P_{acc}^{(M)}(x)$ and $P_{rej}^{(M)}(x)$ denote the accepting and rejecting probabilities in M for input x, respectively.

We say that 2qcfa M recognizes language L over alphabet Σ with one-sided error $\epsilon > 0$ if $P_{acc}^{(M)}(x) = 1$ for $x\in L$, and $P_{rej}^{(M)}(x) \geq 1-\epsilon$ for $x \in L^c = \Sigma^*\backslash L$.

Ambainis and Watrous [2] showed that, for any $\epsilon > 0$, 2qcfa's can recognize palindromes $L_{pal} = \{x \in \{a,b\}^* | x = x^R,\}$ and $L_{eq} = \{a^n b^n | n \in \mathbf{N}\}$ with one-sided error probability ϵ, where ϵ can be arbitrarily small.

Actually, basing on the 2qcfa M_{eq} presented in [2] for recognizing L_{eq} with one-side error, we may further observe that some another non-regular languages can also be recognized by 2qcfa's with bounded error probabilities in polynomial time, and, we would state them in the following Remarks to conclude this section.

. . . . In terms of the 2qcfa M_{eq} by Ambainis and Watrous [2], the language $\{a^n b_1^n a^m b_2^m | n,m \in \mathbf{N}\}$ can also be recognized by some 2qcfa denoted by $M_{eq}^{(2)}$ with one-sided error probability in polynomial time. Indeed, let $M_{eq}^{(2)}$ firstly checks whether or not the input string, say x, is the form $a^{n_1}b_1^{n_2}a^{m_1}b_2^{m_2}$. If not, then x is rejected certainly; otherwise, $M_{eq}^{(2)}$ simulates M_{eq} for deciding whether or not $a^{n_1}b_1^{n_2}$ is in L_{eq}, by using the a in the right of b_1 as the right end-marker \$. If not, then x is rejected; otherwise, this machine continues to simulate M_{eq} for recognizing $a^{m_1}b_2^{m_2}$, in which b_1 is viewed as the left end-marker ¢. If it is accepted, then x is also accepted; otherwise, x is rejected.

. . . For $k \in \mathbf{N}$, let $L_{eq}(k,a) = \{a^{kn}b^n | n \in \mathbf{N}\}$. Obviously, $L_{eq}(1,a) = L_{eq}$. Then, by means of the 2qcfa M_{eq}, $L_{eq}(k,a)$ can be recognized by some 2qcfa, denoted by $M_{eq}(k,a)$, with one-sided error probability in polynomial time. Indeed, $M_{eq}(k,a)$ is derived from M_{eq} by replacing U_β with U_{β_k}, where $\beta_k = \sqrt{2}k\pi$. Likewise, denote $L_{eq}(k,b) = \{b^{kn}a^n | n \in \mathbf{N}\}$. Then $L_{eq}(k,b)$ can be recognized by some 2qcfa $M_{eq}(k,b)$ with one-sided error probability in polynomial time.

. . . Let $L_= = \{x \in \{a,b\}^* | \#_x(a) = \#_x(b)\}$, where $\#_x(a)$ (and $\#_x(b)$) represents the number of a (and b) in string x. Then $L_=$ is recognized by some 2qcfa, denoted by $M_=$, with one-sided error probability in polynomial time.

Indeed, by observing the words in $L_=$, $M_=$ can be directly derived from M_{eq} above by omitting the beginning process for checking whether or not the input string is of the form $a^n b^m$.

3 Operation Properties of 2qcfa's

This section deals with operation properties of 2qcfa's, and, a number of examples as application are incorporated. For convenience, we use notations $2QCFA_\epsilon(poly-time)$ and $2QCFA(poly-time)$ to denote the classes of all languages recognized by 2qcfa's with given error probability $\epsilon \geq 0$ and with unbounded error probabilities in $[0,1)$, respectively, which run in polynomial expected time; for any language $L \in 2QCFA(poly-time)$, let QS_L and CS_L denote respectively the minimum numbers of quantum states and classical states of the 2qcfa that recognizes L with error probability in $[0,1)$. Firstly, we consider intersection operation.

Theorem 1. *If $L_1 \in 2QCFA_{\epsilon_1}(poly-time)$, $L_2 \in 2QCFA_{\epsilon_2}(poly-time)$, then $L_1 \cap L_2 \in 2QCFA_\epsilon(poly-time)$ where $\epsilon = \epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2$*

Proof. See [18]. □

By means of the proof of Theorem 1, we have the following corollaries 1 and 2.

Corollary 1. *Let L_1 and L_2 be languages, and M_1 and M_2 be 2qcfa's with $\epsilon_1, \epsilon_2 \in [0, \frac{1}{2})$, recognizing L_1 and L_2, respectively. Then $L_1 \cap L_2$ is recognized by M with $\epsilon = \max\{\epsilon_1, \epsilon_2\}$, such that, for any input x,*

- *if $x \in L_1 \cap L_2$, then M accepts x;*
- *if $x \notin L_1$, then M accepts x with probability $\leq 1 - \epsilon_1$;*
- *if $x \in L_1$ but $x \notin L_2$, then M accepts x with probability $\leq 1 - \epsilon_2$.*

Example 1. We recall that non-regular language $L_= = \{x \in \{a,b\}^* | \#_x(a) = \#_x(b)\}$. For non-regular language $L_=(pal) = \{y = xx^R | x \in L_=\}$, we can clearly check that $L_=(pal) = L_= \cap L_{pal}$. Therefore, by applying Corollary 1, we obtain that $L_=(pal)$ is recognized by some 2qcfa with one-sided error probability ϵ, since both $L_=$ and L_{pal} are recognized by 2qcfa's with one-sided error probability ϵ [2], where ϵ can be given arbitrarily small.

Corollary 2. *If $L_1 \in 2QCFA(poly-time)$, $L_2 \in 2QCFA(poly-time)$, then $QS_{L_1 \cap L_2} \leq QS_{L_1} + QS_{L_2}$ $CS_{L_1 \cap L_2} \leq CS_{L_1} + CS_{L_2} + QS_{L_1}$*

Similar to Theorem 1, we can obtain the union operation of 2qcfa's.

Theorem 2. *If $L_1 \in 2QCFA_{\epsilon_1}(poly-time)$ and $L_2 \in 2QCFA_{\epsilon_2}(poly-time)$ where $\epsilon_1, \epsilon_2 \geq 0$, then $L_1 \cup L_2 \in 2QCFA_\epsilon(poly-time)$ where $\epsilon = \epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2$*

Proof. See [18]. □

By means of the proof of Theorem 2, we also have the following corollary.

Corollary 3. [illegible] L_1 [illegible] L_2 [illegible] M_1 [illegible] M_2 [illegible] $\epsilon_1, \epsilon_2 \in [0, \frac{1}{2})$ [illegible] M [illegible] $L_1 \cup L_2$ [illegible] M [illegible] $\epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2$ [illegible] x, (a) $x \in L_1$ [illegible] M [illegible] x [illegible] (b) $x \notin L_1$ [illegible] $x \in L_2$ [illegible] M [illegible] x [illegible] $1 - \epsilon_1$ (c) $x \notin L_1$ [illegible] $x \notin L_2$ [illegible] M [illegible] x [illegible] $(1 - \epsilon_1)(1 - \epsilon_2)$

Similar to Corollary 2, we have:

Corollary 4. [illegible] $L_1 \in 2QCFA(poly - time)$, $L_2 \in 2QCFA(poly - time)$, [illegible] (a) $QS_{L_1 \cup L_2} \leq QS_{L_1} + QS_{L_2}$ (b) $CS_{L_1 \cup L_2} \leq CS_{L_1} + CS_{L_2} + QS_{L_1}$

Example 2. As indicated in Remark 2, $L_{eq}(k, a) = \{a^{kn}b^n | n \in \mathbf{N}\}$ and $L_{eq}(k, b) = \{b^n a^n | n \in \mathbf{N}\}$ are recognized by 2qcfa's with one-sided error probabilities (as demonstrated by Ambainis and Watrous [2], these error probabilities can be given arbitrarily small) in polynomial time. Therefore, by using Corollary 3, we have that for any $m \in \mathbf{N}$, $\cup_{k=1}^{m} L_{eq}(k, a)$ and $\cup_{k=1}^{m} L_{eq}(k, b)$ are recognized by 2qcfa's with error probabilities in $[0, \frac{1}{2})$ in polynomial time.

For language L over alphabet Σ, the complement of L is $L^c = \Sigma^* \backslash L$. For the class of languages recognized by 2qcfa's with bounded error probabilities, the unary complement operation is also closed.

Theorem 3. [illegible] $L \in 2QCFA_\epsilon(poly - time)$ [illegible] ϵ [illegible] $L^c \in 2QCFA_\epsilon(poly - time)$

[illegible] See [18]. □

From the proof of Theorem 3 it follows Corollary 5.

Corollary 5. [illegible] $L \in 2QCFA(poly - time)$, [illegible] (a) $QS_{L^c} = QS_L$ (b) $CS_{L^c} = CS_L$

Example 3. For non-regular language $L_=$, its complement $L_=^c = \{x \in \{a, b\}^* | \#_x(a) \neq \#_x(b)\}$ is recognized by 2qcfa with bounded error probability in polynomial expected time, by virtue of Remark 3 and Theorem 3.

For language L over alphabet Σ, the reversal of L is $L^R = \{x^R | x \in L\}$ where x^R is the reversal of x, i.e., if $x = \sigma_1\sigma_2 \ldots \sigma_n$ then $x^R = \sigma_n\sigma_{n-1} \ldots \sigma_1$. For $2QCFA_\epsilon(poly - time)$ with $\epsilon \in [0, 1/2)$, the reversal operation is closed.

Theorem 4. [illegible] $L \in 2QCFA_\epsilon(poly - time)$, [illegible] $L^R \in 2QCFA_\epsilon(poly - time)$

[illegible] See [18]. □

By means of the proof of Theorem 4 we clearly obtain the following corollary.

Corollary 6. [illegible] $L \in 2QCFA(poly - time)$, [illegible] (a) $QS_L - 1 \leq QS_{L^R} \leq QS_L + 1$ (b) $CS_L - 1 \leq CS_{L^R} \leq CS_L + 1$

For languages L_1 and L_2 over alphabets Σ_1 and Σ_2, respectively, the catenation of L_1 and L_2 is $L_1L_2 = \{x_1x_2 | x_1 \in \Sigma_1, x_2 \in \Sigma_2\}$. We do not know whether or not the catenation operation in $2QCFA_\epsilon$ is closed, but under certain condition we can prove that the catenation of two languages in $2QCFA_\epsilon$ is closed.

Theorem 5. [illegible] $L_i \in 2QCFA_\epsilon$ [illegible] $\Sigma_1 \cap \Sigma_2 = \emptyset$ [illegible] Σ_i [illegible] L_i $(i = 1, 2)$ [illegible] L_1L_2 [illegible] L_1 [illegible] L_2 [illegible] $\epsilon = \epsilon_1 + \epsilon_2 - \epsilon_1\epsilon_2$

[illegible] See [18]. □

From Theorem 5 it follows the following corollary.

Corollary 7. [illegible] L_i [illegible] Σ_i [illegible] ϵ_i $(i = 1, 2)$. [illegible] $\Sigma_1 \cap \Sigma_2 = \emptyset$ [illegible] L_1L_2 [illegible] $\max\{\epsilon_1, \epsilon_2\}$ [illegible]

[illegible] As indicated in Remark 1, the catenation $\{a^n b_1^n a^m b_2^m | n, m \in \mathbf{N}\}$ of $L_{eq}^{(1)} = \{a^n b_1^n | n \in \mathbf{N}\}$ and $L_{eq}^{(2)} = \{a^n b_2^n | n \in \mathbf{N}\}$ can also be recognized by some 2qcfa with one-sided error probability ϵ in polynomial time, where ϵ can be arbitrarily small.

4 Concluding Remarks

As a continuation of [2], in this note, we have dealt with a number of operation properties of 2qcfa's. We proved that the Boolean operations (intersection, union, and complement) and the reversal operation of the class of languages recognized by 2qcfa's with error probabilities are closed; as corollaries, we showed that the intersection, complement, and reversal operations in the class of languages recognized by 2qcfa's with one-sided error probabilities (in $[0, \frac{1}{2})$) are closed. Furthermore, we verified that the catenation operation in the class of languages recognized by 2qcfa's with error probabilities is closed under a certain restricted condition (this result also holds for the case of one-sided error probabilities belonging to $[0, \frac{1}{2})$). Also, the numbers of states of these 2qcfa's for the above operations were presented, and some examples were included for an application of the derived results. For instance, $\{xx^R | x \in \{a, b\}^*, \#_x(a) = \#_x(b)\}$ was shown to be recognized by 2qcfa with one-sided error probability $0 \leq \epsilon < \frac{1}{2}$ in polynomial time.

These operation properties presented may apply to 2qfa's [16], but the unitarity should be satisfied in constructing 2qfa's, and, therefore, more technical methods are likely needed or we have to add some restricted conditions (for example, we may restrict the initial state not to be entered again). On the other hand, in Corollaries 2 and 4, the lower bounds need be further fixed. We would like to further consider them in the future.

Acknowledgment

I would like to thank Dr. Tomoyuki Yamakami for helpful discussion regarding qfa's, and thank the anonymous reviewers for their invaluable comments.

References

1. Ambainis, A., Freivalds, R.: One-way quantum finite automata: strengths, weaknesses and generalizations. In: Proc. 39th FOCS, pp. 332–341 (1998)
2. Ambainis, A., Watrous, J.: Two-way finite automata with quantum and classical states. Theoret. Comput. Sci. 287, 299–311 (2002)
3. Benioff, P.: The computer as a physical system: a microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines. J. of Stat.Phys. 22, 563–591 (1980)
4. Bertoni, A., Mereghetti, C., Palano, B.: Quantum Computing: 1-Way Quantum Automata. In: Ésik, Z., Fülöp, Z. (eds.) DLT 2003. LNCS, vol. 2710, pp. 1–20. Springer, Heidelberg (2003)
5. Broadsky, A., Pippenger, N.: Characterizations of 1-way quantum finite automata. SIAM J. Comput. 31, 1456–1478 (2002)
6. Deutsch, D.: Quantum theory, the Church-Turing principle and the universal quantum computer. Proc. R. Soc. Lond. A. 400, 97–117 (1985)
7. Dwork, C., Stockmeyer, L.: A time-complexity gap for two-way probabilistic finite state automata. SIAM J. Comput. 19, 1011–1023 (1990)
8. Dwork, C., Stockmeyer, L.: Finite state verifier I: the power of interaction. J. ACM. 39(4), 800–828 (1992)
9. Feynman, R.P.: Simulating physics with computers. Internat. J.Theoret.Phys. 21, 467–488 (1982)
10. Freivalds, R.: Probabilistic two-way machines. In: Gruska, J., Chytil, M.P. (eds.) MFCS 1981. LNCS, vol. 118, pp. 33–45. Springer, Heidelberg (1981)
11. Grover, L.: A fast quantum mechanical algorithms for datdbase search. In: Proc. 28th STOC, pp. 212–219 (1996)
12. Gruska, J.: Quantum Computing. McGraw-Hill, London (1999)
13. Greenberg, A., Weiss, A.: A lower bound for probabilistic algorithms for finite state machines. J. Comput. System Sci. 33(1), 88–105 (1986)
14. Hopcroft, J.E., Ullman, J.D.: Introduction to Automata Theory, Languages, and Computation. Addison-Wesley, New York (1979)
15. Kaneps, J., Freivalds, R.: Running time to recognize nonregular languages by 2-way probabilistic automata. In: Leach Albert, J., Monien, B., Rodríguez-Artalejo, M. (eds.) ICALP 1991. LNCS, vol. 510, pp. 174–185. Springer, Heidelberg (1991)
16. Kondacs, A., Watrous, J.: On the power of finite state automata. In: Proc. 38th FOCS, pp. 66–75 (1997)
17. Moore, C., Crutchfield, J.P.: Quantum automata and quantum grammars. Theoret. Comput. Sci. 237, 275–306 (2000)
18. Qiu, D.W.: Some observations on two-way finite automata with quantum and classical states, quant-ph/0701187 (2007)
19. Shor, P.W.: Algorithm for quantum computation: discrete logarithms and factoring. In: Proc. 35th FOCS, pp. 124–134 (1994)

Design of DNA Sequence Based on Improved Genetic Algorithm

Bin Wang, Qiang Zhang* and Rui Zhang

Liaoning Key Lab of Intelligent information Processing
Dalian, China
zhangq@dlu.edu.cn

Abstract. DNA computing is a new method that uses biological molecule DNA as computing medium and biochemical reaction as computing tools. The design of DNA sequence plays an important role in improving the reliability of DNA computation. The sequence design is an approach of the control, which aims to design of DNA sequences satisfying some constraints to avoid such unexpected molecular reactions .To deal these constraints ,we convert the design of DNA sequences into a multi-objective optimize problems, and then an example is illustrated to show the efficiency of our method given here.

1 Introduction

In 1994, Dr Adleman released "Molecular Computation of Solutions to Combinatorial Problems" in Science, which indicates a new research field-DNA computing comes into being [1]. DNA computing is a new method that uses biological molecule DNA as computing medium and biochemical reaction as computing tools [2]. In DNA computation, the core reaction is the specific hybridization between DNA sequences or the Watson-Crick complement, and which directly influences the reliability of DNA computation with its efficiency and accuracy. However, false hybridization can also occur [3].False hybridization in DNA computation process can be assorted two categories [4, 5]. One is false positive, the other is false negative.

The encoding problem means trying to encode every bit of DNA code words in order to make the DNA strands hybridize with its complement specifically in biochemical reaction. Presently, the goal of encoding is mainly to minimize the similarity distance of the various DNA sequences by using some convenient similarity measure. So as a necessary criterion for relia[1]ble DNA computation, the minimum Hamming distance [3] and H-measure based on Hamming distance [6] between DNA code words were proposed to define the distance. And various algorithms and methods for the reliable DNA sequence design were presented based on the two distance measures. For example, Deaton et al. proposed an evolution search method [7].

In the previous works, the encoding problems were considered as a numeral optimization problem that satisfying constraints based on knowledge of sequence design. Among the optimization methods, the GA Algorithm is probably the most popular method of parameter optimization for a problem which is difficult to be

*Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 9–14, 2008.
© Springer-Verlag Berlin Heidelberg 2008

mathematically formalized. In Soo-Yong shin's paper [13], better sequences were generated by using GA Algorithm to solve the multi-objective optimization problem. However, in the previous works, the more constraints are, the lower values of fitness function are. Further more, the larger number of DNA sequences are, the lower values of fitness function are. At the same time, the inherent defaults, along whit convergent in local values in GA algorithm, were not considered. Corresponding to these observations, in this paper, we use an improved GA to solve the multi-objective problem on the constraints of DNA sequence design.

2 DNA Encoding

2.1 DNA Encoding Criterions

The encoding problem can be described as follows: the encoding alphabet of DNA sequence is the set of nucleic acid bases A, T, G, C, and in the encoding set S of DNA strands whose length is N, search the subset C of S which satisfies for $\forall x_i, x_j \in C$, $\tau(x_i, x_j) \geq k$, where k is a positive integer, and τ is the expected criterion of evaluating the encoding, namely the encodings should satisfy constraint conditions. Theoretically, the encoding should satisfy two kinds of constraints: combinatorial constraints and thermodynamic constraints [12].

2.2 DNA Encoding Constraints [8] [9] [10] [11] [12] [13]

Hamming Distance

Hamming constraint: the Hamming distance between x_i and x_j should not be less than certain parameter d , i.e. $H(x_i, x_j) \geq d$. Where $f'_{Ham\min g}(i)$ indicates the Hamming evaluation function of the i th individual in evolutionary population.

$$f'_{Hm\min g}(i) = \min_{1 \leq j \prec m,\, j \neq i} \{H(x_i, x_j)\}. \tag{1}$$

Reverse Hamming Constraint

Reverse constraint: the Hamming distance between x_i and x_j^R should not be less than certain parameter d, i.e. $H(x_i, x_j^R) \geq d$,

$$f'_{\mathrm{Re}verse}(i) = \min_{1 \leq j \prec m} \{H(x_i, x_j^R)\}. \tag{2}$$

Reverse-Complement Hamming Constraint

Reverse-complement hamming constraint: the Hamming distance between x_i' and x_j^R should not be less than certain parameter d, i.e. $H(x_i', x_j^R) \geq d$,

$$f'_{\mathrm{R}everse_Comple}(i) = \min_{1 \leq j \prec m} \{H(x_i', x_j^R)\}. \tag{3}$$

Continuity Constraint
If the same base appears continuously, the structure of DNA will become unstable. The evaluation function is described as follow:

$$f_{\mathrm{Con}}(i) = -\sum_{j=1}^{n} (j-1)\,\mathrm{N}_j^{(i)} \ . \tag{4}$$

where $\mathrm{N}_j^{(i)}$ denotes the number of time to which the same base appears j-times continuously in sequence x_i .

GC Content Constraint
GC content affects the chemical properties of DNA sequence.

$$f_{GC}(i) = -\lambda \left| GC^{(i)} - GC_{defined}^{(i)} \right| \ . \tag{5}$$

where $GC_{defined}^{(i)}$ is the target value of the GC content of DNA sequence x_i , and $GC^{(i)}$ is the real GC content, λ is the parameter that is used to adjust the weight of the constraint and other constraints.

Fitness function
Based on the design of DNA sequences combining with various improved conventional constraint terms, and is transformed into a multi-objective optimization problem. The fitness function is the weighted sum of the required fitness measures:

$$f_t(i) \in \left\{ \begin{matrix} f_{Ham\min g}(i),\ f_{\mathrm{R}everse}(i),\ f_{\mathrm{R}everse-Comple}(i), \\ f_{\mathrm{Con}}(i),\ f_{GC}(i) \end{matrix} \right\}$$

$$f(i) = \sum_{j=1}^{5} \omega_j f_j(i) \ . \tag{6}$$

where ω_j is the weight of the each constraint term.

3 Design of Algorithm

In this paper, we use an improved GA to solve the multi-objective optimization problem. The main idea is that using the Best-Saved strategy to improved GA algorithm. This method can improve the speed of convergence of GA algorithm and save best populations. Then it could avoid destroying the best population. In the best populations, we delete the least value of fitness of best population, and it is joined the new population. A new best population is joined in the best populations, if the number of best populations is not enough and the least value of fitness of best population is less than 32.

The steps of improved GA algorithm solving the sequence design are as follows:

Step1: Set parameter, and initialize population randomly.
Step2: Calculate the fitness value of every individual in populations.
Step3: Select the most fitness value of population and join it in the best population.

Step4: Generate next population by select、crossover and mutate. If the number of best population is not enough, go to Step 2. Otherwise go to Step 5.
Step5: Calculate the fitness value of every individual in best populations. If the least value of fitness of best population is less than 32, delete this best population and go to Step4.Otherwise, go to step6.
Step6: end.

4 Simulate Results

The improved GA mentioned above is implemented with Matlab 7.1. The parameters of improved GA used in our example are: the size of population is 8, which is equal to the size of best population. The length of DNA sequences is 20. The probability of crossover is 0.6. The probability of mutate is 0.01.

Fig.1 illustrates the results of simulation, where the x-axis is the generation of population, and y-axis is the lowest fitness value among the best populations. Before the 46 generation, the fitness is 27. After the 275 generation, the fitness is 33. We performed fifty trials, and there are thirty-four times that the number of generation is under three hundred.

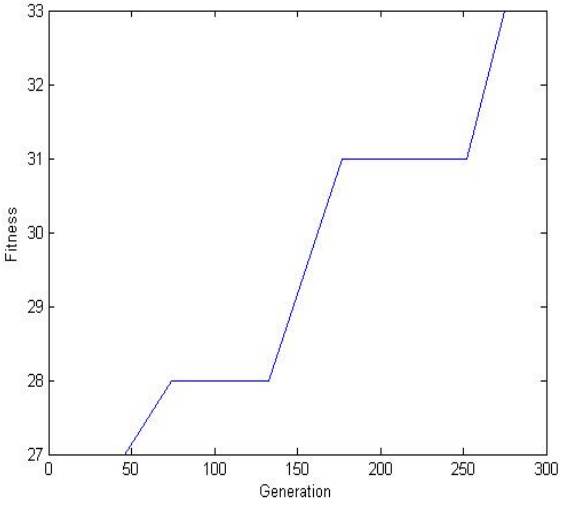


Fig. 1. Results of simulation

5 Comparison in DNA Sequence

To evaluate the performance of algorithms, eight good sequences are generated, and are compared with the sequences in the previous works under the various rules.

Table 1 shows the value of DNA sequence based on Hamming Distance, Reverse Hamming, Reverse-Complement Hamming, Continuity, GC Content. In our example,

Table 1. Eight good DNA sequences in our system and seven in Soo-Yong Shin's paper

DNA sequences in our system						
DNA Sequence(5 ' - 3 ')	$f_{Hamming}$	$f_{Reverse}$	$f_{Reverse-Complement}$	f_{GC}	f_{Con}	Fitness
GTCGAAACCTGAGGTACAGA	12	11	14	50	-4	33
GTGACTGTATGCACTCGAGA	12	14	12	50	0	38
ACGTAGCTCGAATCAGCACT	12	11	14	50	-1	36
CGCTGATCTCAGTGCGTATA	11	11	14	50	0	36
GTCAGCCAATACGAGAGCTA	12	13	14	50	-2	37
GCATGTCTAACTCAGTCGTC	12	11	13	50	-1	35
TACTGACTCGAGTAG-GACTC	11	13	12	50	-1	35
GGTTCGAGTGTTCCAAGTAG	15	11	13	50	-5	34
Sequences Soo-Yong Shin's paper[13]						
CGCTCCATCCTTGATCGTTT	11	13	13	50	-5	33
CCTGTCAACATT-GACGCTCA	11	11	14	50	-3	33
CTTCGCTGCTGATAACCTCA	11	10	11	50	-3	29
TTATGATTCCACTGGCGCTC	13	13	11	50	-4	33
GAGTTAGATGTCACGTCACG	15	14	13	50	-1	41
AGGCGAGTATGGGGTATATC	14	12	13	50	-4	35
ATCGTACTCATGGTCCCTAC	11	10	12	50	-3	30

larger and better fitness value of DNA sequences are obtained by comparing with DNA sequence in Soo-Yong Shin's paper [13].

6 Conclusions

In this paper, some constraint terms based on Hamming Distance are selected from various constraint terms, and then the selected terms are transformed to a multi-objective optimization problem. The Improved Genetic Algorithm is proposed to solve the optimization problem, and good sequences are obtained to improve the reliability of DNA computation. The feasibility and the efficiency of our system are proved by comparing our sequence with other sequences. However this paper does not consider any thermodynamic factors (i.e. melting temperature). Therefore, this paper needs more study.

Acknowledgments

This paper was supported by the National Natural Science Foundation of China (Grant Nos. 60403001, 60533010 and 30740036).

References

1. Adleman, L.: Molecular Computation of Solution to Combinatorial problems. Science, 1021–1024 (1994)
2. Cui, G.Z., Niu, Y.Y., Zhang, X.C.: Progress in the Research of DNA Codeword Design. Biotechnology Bulletin, 16–19 (2006)
3. Deaton, R., Murphy, R.C., Garzon, M., Franceshetti, D.R., Stevens Jr., S.E.: Good Encodings for DNA-based Solutions to Combinatorial Problems. In: Proceedings of 2nd DIMACS Workshop on DNA Based Computers, pp. 159–171 (1996)
4. Deaton, R., Garzon, M.: Thermodynamic Constraints on DNA-based Computing. In: Paun, G. (ed.) Computing with Bio-Molecus: Theory and Experiments, pp. 138–152. Springer, Singapore (1998)
5. Deaton, R., Franceschetti, D.R., Garzon, M., Rose, J.A., Murphy, R.C., Stevens, S.E.: Information Transfer Through Hybridization Reaction in DNA based Computing. In: Proceedings of the Second Annual Conference, Stanford University, vol. 13, pp. 463–471. Morgan Kaufmann, San Francisco (1997)
6. Garzon, M., Neathery, P., Deaton, R.J., Murphy, R.C., Franceschetti, D.R., Stevens Jr., S.E.: A new Metric for DNA Computing. In: Proceedings of the 2nd Annual Gentic Programming Conference, pp. 472–487 (1997)
7. Deaton, R., Murphy, R.C., Rose, J.A., Garzon, M., Franceschett, D.R., Stevens, S.E.: A DNA Based Implementation of an Evolutionary Search for Good Encodings for DNA Computation. In: Proceedings of the 1997 IEEE International Conference on Evolutionary Computation, pp. 267–272 (1997)
8. Marathe, A., Condon, A.E., Corn, R.M.: On combinatorial DNA Word Design. In: Proceedings of 5th DIMACS Workshop on DNA Based Computers, pp. 75–89 (1999)
9. Arita, M., Nishikawa, A., Hagiya, M., Komiya, K., Gouzu, H., Sakamoto, K.: Improving Sequence Design for DNA Computing. In: Proceedings of Genetic and Evolutionary Computation Conference 2000, pp. 875–882 (2000)
10. Tanaka, F., Nakatsugawa, M., Yamamoto, M., Shiba, T., Ohuchi, A.: Developing Support System for Sequence Design in DNA Computing. In: Preliminary Proceedings of 7th international Workshop on DNA-Based Computers, pp. 340–349 (2001)
11. Deaton, R., Garzon, M., Murphy, R.R.J., Franceschetti, D., Stevens Jr., S.: Reliability and Efficiency of a DNA-Based Computation. Physical Review Letters, 417–420 (1998)
12. Wang, W., Zheng, X.D., Zhang, Q., Xu, J.: The Optimization of DNA Encodings Based on GA/SA Algorithms. Progress in Natural Science 17, 739–744 (2007)
13. Shin, S.Y., Kim, D.M., Zhang, B.T.: Evolutionary sequence generation for reliable DNA computing. Evolutionary Computation (2002)

Bijective Digital Error-Control Coding, Part I: The Reversible Viterbi Algorithm

Anas N. Al-Rabadi

Computer Engineering Department, The University of Jordan, Jordan
alrabadi@yahoo.com
http://web.pdx.edu/~psu21829/

Abstract. New convolution-based multiple-stream digital error-control coding and decoding schemes are introduced. The new digital coding method applies the reversibility property in the convolution-based encoder for multiple-stream error-control encoding and implements the reversibility property in the new reversible Viterbi decoding algorithm for multiple-stream error-correction decoding. It is also shown in this paper that the reversibility relationship between multiple-streams of data can be used for further correction of errors that are uncorrectable using the implemented decoding algorithm such as in triple-errors that are uncorrectable using the classical (irreversible) Viterbi algorithm.

Keywords: Error-Correcting Codes, Quantum Computing, Reversible Logic.

1 Introduction

Due to the anticipated failure of Moore's law around the year 2020, quantum computing will play an increasingly crucial role in building more compact and less power consuming computers [1,13]. Due to this fact, and because all quantum computer gates (i.e., building blocks) should be reversible [1,3,7,12,13], reversible computing will have an increasingly more existence in the future design of regular, compact, and universal circuits and systems. It is shown in [12] that the amount of energy (heat) dissipated for every irreversible bit operation is given by $K \cdot T \cdot ln2$ where K is the Boltzmann constant and T is the operating temperature, and that a necessary (but not sufficient) condition for not dissipating power in any physical circuit is that all system circuits must be built using fully reversible logical components.

In general, in data communications between two communicating systems, noise exists and corrupts sent data messages, and thus noisy corrupted messages will be received [5,10,11]. The corrupting noise is usually sourced from the communication channel. Many solutions have been classically implemented to solve for the classical error detection and correction problems: (1) one solution to solve for error-control is *parity checking* [4,5,11,15] which is one of the most widely used methods for error detection in digital logic circuits and systems, in which re-sending data is performed in case error is detected in the transmitted data. Various parity-preserving circuits have been implemented in which the parity of the outputs matches that of the inputs, and such circuits can be fault-tolerant since a circuit output can detect a single error;

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 15 – 22, 2008.
© Springer-Verlag Berlin Heidelberg 2008

(2) another solution to solve this highly important problem, that is to extract the correct data message from the noisy erroneous counterpart, is by using various coding schemes that work optimally for specific types of statistical distributions of noise [5,8,9,17]. For example, the manufacturers of integrated circuits (ICs) have recently started to produce error-correcting circuits, and one such circuit is the TI 74LS636 [2] which is an 8-bit error detection and correction circuit that corrects any single-bit memory read error and flags any two-bit error which is called single error correction / double error detection (SECDED). Figure 1 shows the 74LS636 IC that corrects errors by storing five parity bits with each byte of memory data [2].

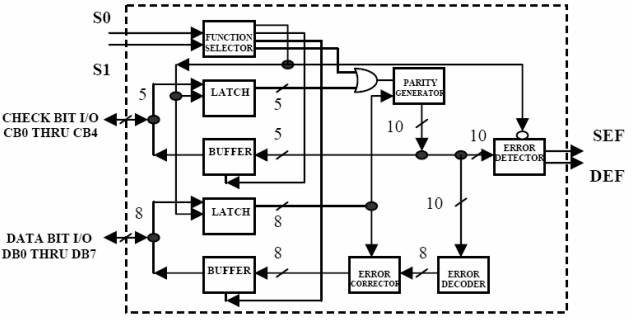


Fig. 1. Block diagram of the Texas Instruments (TI) 74LS636 error-correcting IC

2 Fundamentals

This Section presents basic background in the topics of error-correction coding and reversible logic that will be utilized in the the new results introduced in Section 3.

2.1 Error Correction

In the data communication context, noise usually exists and is generated from the channel in which transmitted data are communicated. Such noise corrupts sent messages from one end and thus noisy corrupted messages are received on the other end. To solve the problem of extracting a correct message from its corrupted counterpart, noise must be modeled [5,6,17] and accordingly an appropriate encoding / decoding communication schemes must be implemented [4,5,6,8,9,17]. Various coding schemes have been proposed and one very important family is the convolutional codes. Figure 2 illustrates the modeling of data communication in the existence of noise, the solution to the noise problem using an encoder / decoder scheme, and the utilization of a new block called the “reverser” for bijectivity (uniqueness) in multiple-stream (i.e., parallel data) communication.

Definition 1. For an L-bit message sequence, M-stage shift register, n mouulo-2 adders, and a generated coded output sequence of length $n(L + M)$ bits, the code rate r is calculated as: $r = \dfrac{L}{n(L+M)}$ bits / symbol. (Usually $L>>M$.)

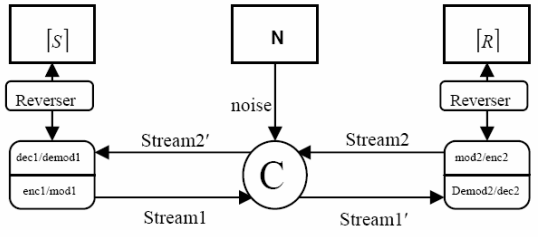


Fig. 2. Model of a noisy data communication using the application of reversibility using the reverser block for parallel multiple-input multiple-output (MIMO) bijectivity in data streams

Definition 2. The constraint length of a convolutional code is the number of shifts over which a single message bit can influence the encoder output.

Each path connecting the output to the input of a convolutional encoder can be characterized in terms of the impulse response which is defined as the response of that path to "1" applied to its input, with each flip-flop of the encoder set initially to "0". Equivalently, we can characterize each path in terms of a generator polynomial defined as the unit-delay transform of the impulse response. More specifically, the generator polynomial is defined as:

$$g(D) = \sum_{i=0}^{M} g_i D^i \ , \tag{1}$$

where g_i is the generator coefficients $\in \{0, 1\}$, and the generator sequence $\{g_0, g_1, \ldots, g_M\}$ composed of generator coefficients is the impulse response of the corresponding path in the convolutional encoder, and D is the unit-delay variable.

The functional properties of the convolutional encoder (e.g., Figure 3) can be represented graphically as a trellis (cf. Figure 4). An important decoder that uses the trellis representation to correct received erroneous messages is the Viterbi decoding algorithm [9,17]. The Viterbi algorithm is a dynamic programming algorithm which is used to find the maximum-likelihood sequence of hidden states, which results in a sequence of observed events particularly in the context of hidden Markov models (HMMs) [14]. The Viterbi algorithm forms a subset of information theory [6], and has been extensively used in a wide range of applications such as in wireless communications and networks.

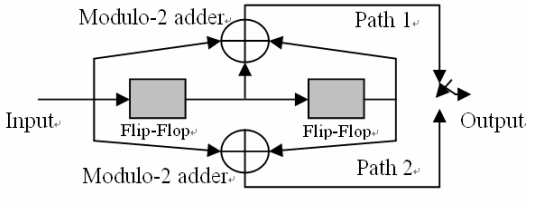


Fig. 3. Convolutional encoder with constraint length = 3 and rate = ½ . The modulo-2 adder is the logic Boolean difference (XOR).

The Viterbi algorithm is a maximum-likelihood decoder (i.e., maximum-likelihood sequence estimator) which is optimum for a noise type which is statistically characterized as an Additive White Gaussian Noise (AWGN). The following procedure illustrates the steps of this algorithm.

Algorithm. Viterbi

1. Initialization: Label the left-most state of the trellis (i.e., all zero state at level 0) as 0
2. Computation step $j + 1$: Let $j = 0, 1, 2, \ldots$, and suppose at the previous j the following is done:
   a. All survivor paths are identified
   b. The survivor paths and its metric for each state of the trellis are stored

   Then, at level (clock time) j +1, compute the metric for all the paths entering each state of the trellis by adding the metric of the incoming branches to the metric of the connecting survivor path from level j. Thus, for each state, identify the path with the lowest metric as the survivor of step $j + 1$, therefore updating the computation
3. Final step: Continue the computation until the algorithm completes the forward search through the trellis and thus reaches the terminating node (i.e., all zero state), at which time it makes a decision on the maximum-likelihood path. Then, the sequence of symbols associated with that path is released to the destination as the decoded version of the received sequence

2.2 Reversible Logic

In general, an (n, k) reversible circuit is a circuit that has n number of inputs and k number of outputs and is one-to-one mapping between vectors of inputs and outputs, thus the vector of input states can be always uniquely reconstructed from the vector of output states [1,3,7,12,13,16]. Thus, a (k, k) *reversible* map is a *bijective* function which is both (1) *injective* and (2) *surjective*. The *auxiliary* outputs that are needed *only* for the purpose of reversibility are called "garbage" outputs. These are auxiliary outputs from which a reversible map is constructed. Therefore, reversible circuits (systems) are *information-lossless*. Geometrically, achieving reversibility leads to value space-partitioning that leads to spatial partitions of unique values.

Algebraically, and in terms of systems representation, reversibility leads to multi-input multi-output (MIMO) bijective maps (i.e., bijective functions). One form of an algorithm which is called reversible Boolean function (RevBF) [1] that produces a reversible map from an irreversible Boolean function will be used in the following section for the new method of bijective coding.

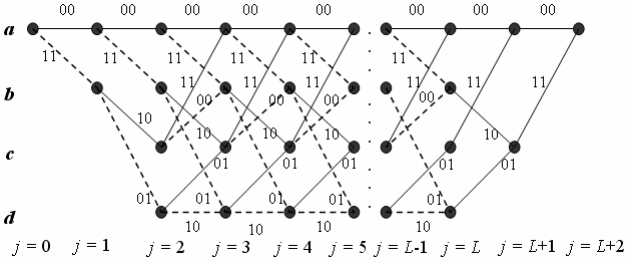


Fig. 4. Trellis form for the convolutional encoder in Fig. 3. Solid line is the input of value "0" and the dashed line is the input of value "1".

3 Bijective Error Correction Via Reversible Viterbi Algorithm

While in subsection 2.1 error correction of communicated data was performed for the case of single-input single-output (SISO) systems, this section introduces reversible error correction of communicated batch (parallel) of data in multiple-input multiple-output (MIMO) systems. Reversibility in parallel-based data communication is directly observed since:

$$[O_1] = [I_2], \tag{2}$$

where $[O_1]$ is the *unique* output (transmitted) data from node1 and $[I_2]$ is the *unique* input (received) data to node2.

In MIMO systems, the existence of noise will cause an error that may lead to irreversibility in data communication (i.e., irreversibility in data mapping) since $[O_1] \neq [I_2]$. The following new algorithm, called Reversible Viterbi (RV) Algorithm, introduces the implementation of reversible error correction in the communicated parallel data.

Algorithm. RV

1. Use the RevBF Algorithm to reversibly encode the communicated batch of data
2. Given a specific convolutional encoder circuit, determine the generator polynomials for all paths
3. **For** each communicated message within the batch, determine the encoded message sequence
4. **For** each received message, use the Viterbi Algorithm to decode the received erroneous message
5. Generate the total maximum-likelihood trellis resulting from the *iterative* application of the Viterbi decoding algorithm
6. Generate the corrected communicated batch of data messages
7. **End**

The convolutional encoding for the RV algorithm can be performed in *parallel* using a general parallel convolutional encoder circuit in which several *s* convolutional encoders operate in parallel for encoding *s* number of simultaneously submitted messages generated from *s* nodes. This is a general MIMO encoder circuit for the parallel generation of convolutional codes where each box represents a single SISO convolutional encoder such as the one shown in Figure 3.

Example 1. The reversibility implementation (cf. RevBF Algorithm) upon the following input bit stream $\{m_1 = 1, m_2 = 1, m_3 = 1\}$ produces the following reversible set of message sequences: $m_1 = (101)$, $m_2 = (001)$, $m_3 = (011)$. For the MIMO convolutional encoder, the following is the D-domain polynomial representations, respectively: $m_1(D) = 1 \cdot D^0 + 0 \cdot D^1 + 1 \cdot D^2 = 1 + D^2$, $m_2(D) = 0 \cdot D^0 + 0 \cdot D^1 + 1 \cdot D^2 = D^2$, $m_3(D) = 0 \cdot D^0 + 1 \cdot D^1 + 1 \cdot D^2 = D + D^2$. The resulting encoded sequences are generated in parallel as follows: $c_1 = (1110001011)$, $c_2 = (0000111011)$, $c_3 = (0011010111)$. Now suppose noise sources corrupt these sequences as follows: $c'_1 = (1111001001)$, $c'_2 = (0100101011)$, $c'_3 = (0010011111)$. Using the RV algorithm, Figure 5 shows the survivor paths which generate the correct sent messages: $\{c_1 = (1110001011), c_2 = (0000111011), c_3 = (0011010111)\}$.

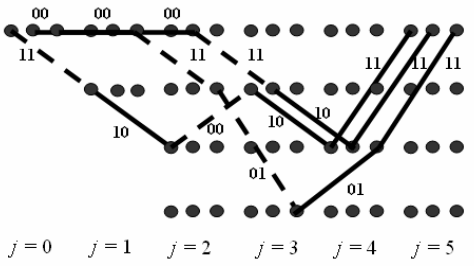


Fig. 5. The resulting survivor paths of the RV algorithm when applied to Example 1

As in the irreversible Viterbi Algorithm, in some cases, applying the reversible Viterbi (RV) algorithm leads to the following difficulties: (1) when the paths entering a node (state) are compared and their metrics are found to be identical then a choice is made by making a guess; (2) when the received sequence is very long and in this case the reversible Viterbi algorithm is applied to a truncated path memory using a decoding window of length $\geq 5K$; (3) the number of errors > 2.

Yet, parallelism in multi-stream data transmission allows for the possible existence of extra relationship(s) between the submitted data-streams that can be used for (1) detection of error existence and (2) further correction after RV algorithm in case the RV algorithm fails to correct for the occurring errors. Examples of such inter-stream relationships are: (1) parity, (2) reversibility, or (3) combination of parity and reversibility properties. *The reversibility property in the RV algorithm produces a reversibility relationship between the sent parallel streams of data, and this known reversibility mapping can be used to correct the uncorrectable errors (e.g., triple errors) which the RV algorithm fails to correct.*

Example 2. The following is RevBF V1 algorithm that produces reversibility:

Algorithm. RevBF (Version 1)

1. To achieve (k, k) reversibility, add sufficient number of auxiliary output variables (starting from right to left) such that the number of outputs equals the number of inputs. Allocate a new column in the mapping's table for each auxiliary variable
2. For construction of the first auxiliary output, assign a constant C_1 = "0" to half of the cells in the corresponding table column, and the second half as another constant C_2 = "1". Assign C_1 to the first half of the column, and C_2 to the second half of the column
3. For the next auxiliary output, **If** non-reversibility still exists, **Then** assign for *identical* output tuples (irreversible map entries) values which are half ones and half zeros, and then assign a constant for the remainder that are already reversible which is the one's complement (NOT; inversion) of the previously assigned constant to that remainder
4. **Do** step 3 until all map entries are reversible

For the parallel sent bit stream {1,1,1} in Example 1 in which the reversibility implementation (using RevBF V1 Algorithm) produces the following reversible sent set of data sequences: $\{m_1 = (101), m_2 = (001), m_3 = (011)\}$. Suppose that m_1 and m_2 are decoded correctly and m_3 is still erroneous due to submission. Figure 6 shows possible tables in which an erroneous m_3 exists.

Note that the erroneous m_3 is Figures 6b-6e and 6g-6h are correctable using the RV algorithm since less than triple-errors exits, but triple errors as in Figure 6f are (usually) uncorrectable using the RV algorithm. Yet, the existence of the reversibility property using the RevBF algorithm adds information that can be used to correct m_3 as follows: By applying the RevBF Algorithm (Version 1) from right-to-left in Figure 6f one notes that in the second column (from right) two "0" cells are added in the top in the correctly received m_1 and m_2 messages, which means that in the most right column the last cell must be "1" since otherwise the top two cells in the correctly received m_1 and m_2 messages should have been "0" and "1" respectively to achieve unique value space-partitioning. Now, since the 3rd cell of the most right column must be "1" then the last cell of the 2nd column from the right must be "1" also because of the uniqueness requirement according to the RevBF algorithm (Version 1) for unique value space-partitioning between the first two messages $\{m_1, m_2\}$ and the 3rd message m_3. Then, and according to the RevBF algorithm (Version 1) the 3rd cell of the last column from right must have the value "0" which is the one's complement (NOT) of the previously assigned constant "1" to the 3rd cell of the 2nd column from the right. Consequently, the correct message $m_3 = (011)$ is obtained.

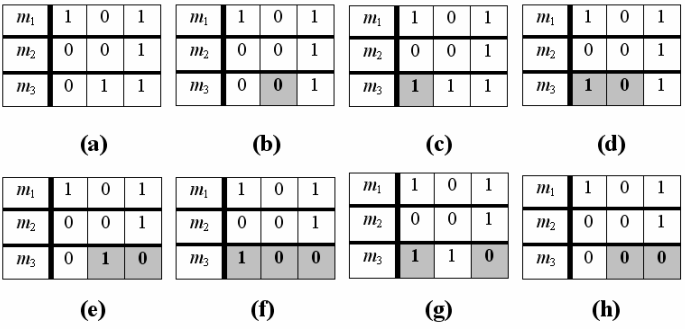


Fig. 6. Tables for possible errors in data stream m_3 that is generated by the RevBF Algorithm V1: (a) original sent uncorrupted m_3 that resulted from the application of the RevBF V1 Algorithm, and (b) – (h) possibilities of the erroneous received m_3

4 Conclusions

This paper introduces new convolution-based multiple-stream error-correction encoding and decoding methods that implement the reversibility property in the convolution-based encoder for multiple-stream error-control encoding and in the new reversible Viterbi (RV) decoding algorithm for multiple-stream error-control decoding. It is also shown in this paper that the relationship of reversibility in multiple-streams of communicated parallel data can be used for further correction of errors that are uncorrectable using the implemented decoding algorithm such as in the cases of the failure of the RV algorithm in correcting for more than two errors.

References

1. Al-Rabadi, A.N.: Reversible Logic Synthesis: From Fundamentals to Quantum Computing. Springer, New York (2004)
2. Brey, B.B.: The Intel Microprocessors, 7th edn. Pearson Education, London (2006)
3. Bennett, C.: Logical Reversibility of Computation. IBM J. Res. Dev. 17, 525–532 (1973)
4. Berlekamp, E.R.: Algebraic Coding Theory. McGraw-Hill, New York (1968)
5. Clark Jr., G.C., Cain, J.B.: Error-Correction Coding for Digital Communications. Plenum Publishers, New York (1981)
6. Cover, T., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)
7. De Vos, A.: Reversible Computing. Progress in Quant. Elect. 23, 1–49 (1999)
8. Divsalar, D.: Turbo Codes. MILCOM tutorial, San Diego (1996)
9. Forney, G.D.: The Viterbi Algorithm. Proc. of the IEEE 61(3), 268–278 (1973)
10. Gabor, D.: Theory of Communications. J. of lEE 93, Part ifi, 429–457 (1946)
11. Gallager, R.G.: Low-Density Parity-Check Codes. MIT Press, Cambridge (1963)
12. Landauer, R.: Irreversibility and Heat Generation in the Computational Process. IBM J. Res. and Dev. 5, 183–191 (1961)
13. Nielsen, M., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)
14. Rabiner, L.R.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE 77(2), 257–286 (1989)
15. Reed, I.S., Solomon, G.: Polynomial Codes Over Certain Finite Fields. Journal of SIAM 8, 300–304 (1960)
16. Roy, K., Prasad, S.: Low-Power CMOS VLSI Circuit Design. John Wiley, Chichester (2000)
17. Viterbi, A.J.: Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm. IEEE Trans. Inf. Theory. IT 13, 260–269 (1967)

Bijective Digital Error-Control Coding, Part II: Quantum Viterbi Circuit Synthesis

Anas N. Al-Rabadi

Computer Engineering Department, The University of Jordan, Jordan
alrabadi@yahoo.com
http://web.pdx.edu/~psu21829/

Abstract. The complete design of quantum circuits for the quantum realization of the new quantum Viterbi cell in the quantum domain (Q-domain) is introduced. Quantum error-control coding can be important for the following main reasons: (1) low-power fault-correction circuit design in future technologies such as in quantum computing (QC), and (2) super-speedy encoding/decoding operations because of the superposition and entanglement properties that emerge in the quantum computing systems and therefore very high performance is obtained.

Keywords: Error-Control Coding, Quantum Circuits, Reversible Circuits.

1 Introduction

Motivations for pursuing the possibility of implementing circuits and systems using quantum computing (QC) would include items such as: (1) power: the fact that, theoretically, the internal computations in QC systems consume no power; (2) size: the current trends related to more dense hardware implementations are heading towards 1 Angstrom (atomic size), at which quantum mechanical effects have to be accounted for; and (3) speed (performance): if the properties of superposition and entanglement of quantum mechanics can be usefully employed in the systems design, significant computational speed enhancements can be expected [1,3].

In general, in data communications between two communicating systems, noise exists and corrupts sent data messages. Many solutions have been classically implemented to solve for the classical error detection and correction problems: (1) one solution to solve for error-control is parity checking which is one of the most widely used methods for error detection in digital logic circuits and systems, in which resending data is performed in case error is detected in the transmitted data; (2) another solution to solve this important problem, that is to extract the correct data message from the noisy erroneous counterpart, is by using various coding schemes that work optimally for specific types of statistical distributions of noise [4].

2 Fundamentals

This Section presents basic background in the topics of error-correction coding and quantum computing that will be utilized in the new results introduced in Section 3.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 23–30, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2.1 Error Correction

In data communication, noise usually occurs and is mostly generated from the channel in which transmitted data are communicated. Such noise corrupts sent messages from one end and thus noisy erroneous messages are received on the other end. To solve the problem of extracting a correct message from its erroneous counterpart, noise must be modeled and accordingly an appropriate encoding / decoding communication schemes must be implemented [4]. Various coding methods have been proposed and one important family is the convolutional codes [4].

An important decoder that uses the trellis representation to correct received erroneous messages is the Viterbi decoding algorithm [4]. The Viterbi algorithm is a dynamic programming algorithm which is used to find the maximum-likelihood sequence of hidden states, which results in a sequence of observed events particularly in the context of hidden Markov models (HMMs). The Viterbi algorithm forms a subset of information theory, and has been extensively used in a wide range of applications. The Viterbi algorithm is a maximum-likelihood decoder (which is a maximum-likelihood sequence estimator) and is optimum for a noise type which is statistically characterized as an Additive White Gaussian Noise (AWGN).

2.2 Quantum Computing

Quantum computing (QC) is a method of computation that uses a dynamic process governed by the Schrödinger Equation (SE) [1,3]. The one-dimensional time-dependent SE (TDSE) takes the following general forms ((1) or (2)):

$$-\frac{(h/2\pi)^2}{2m}\frac{\partial^2|\psi\rangle}{\partial x^2}+V|\psi\rangle=i(h/2\pi)\frac{\partial|\psi\rangle}{\partial t} \tag{1}$$

$$H|\psi\rangle=i(h/2\pi)\frac{\partial|\psi\rangle}{\partial t} \tag{2}$$

where h is Planck's constant ($6.626\cdot10^{-34}$ J·s), $V(x,t)$ is the potential, m is particle's mass, i is the imaginary number, $|\psi(x,t)\rangle$ is the quantum state, H is the Hamiltonian operator ($H = -[(h/2\pi)^2/2m]\nabla^2 + V$), and ∇^2 is the Laplacian operator. While the above holds for all physical systems, in the quantum computing (QC) context, the time-independent SE (TISE) is normally used [1,3]:

$$\nabla^2\psi=\frac{2m}{(h/2\pi)^2}(V-E)\psi \tag{3}$$

where the solution $|\psi\rangle$ is an expansion over orthogonal basis states $|\phi_i\rangle$ defined in Hilbert space $\mathbf{H}$ as follows:

$$|\psi\rangle=\sum_i c_i|\phi_i\rangle \tag{4}$$

where the coefficients c_i are called *probability amplitudes*, and $|c_i|^2$ is the probability that the quantum state $|\psi\rangle$ will collapse into the (eigen) state $|\phi_i\rangle$. The probability is

equal to the inner product $\left|\langle \phi_i \mid \psi \rangle\right|^2$, with the unitary condition $\sum|c_i|^2 = 1$. In QC, a linear and unitary operator $\mathfrak{I}$ is used to transform an input vector of quantum bits (qubits) into an output vector of qubits [1,3]. In two-valued QC, a qubit is a vector of bits defined as follows:

$$qubit_0 \equiv |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, qubit_1 \equiv |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{5}$$

A two-valued quantum state $|\psi\rangle$ is a superposition of quantum basis states $|\phi_i\rangle$ such as those defined in Equation (5). Thus, for the orthonormal computational basis states $\{|0\rangle, |1\rangle\}$, one has the following quantum state:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{6}$$

where $\alpha\alpha^* = |\alpha|^2 = p_0 \equiv$ the probability of having state $|\psi\rangle$ in state $|0\rangle$, $\beta\beta^* = |\beta|^2 = p_1 \equiv$ the probability of having state $|\psi\rangle$ in state $|1\rangle$, and $|\alpha|^2 + |\beta|^2 = 1$. The calculation in QC for multiple systems (e.g., the equivalent of a register) follow the tensor product ($\otimes$) [1,3]. For example, given two states $|\psi_1\rangle$ and $|\psi_2\rangle$ one has the following QC:

$$\begin{aligned} |\psi_{12}\rangle &= |\psi_1\psi_2\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \\ &= \alpha_1\alpha_2|00\rangle + \alpha_1\beta_2|01\rangle + \beta_1\alpha_2|10\rangle + \beta_1\beta_2|11\rangle \end{aligned} \tag{7}$$

A physical system, describable by the following equation [1,3]:

$$|\psi\rangle = c_1|Spinup\rangle + c_2|Spindown\rangle \tag{8}$$

(e.g., the hydrogen atom), can be used to physically implement a two-valued QC. Another common alternative form of Equation (8) is:

$$|\psi\rangle = c_1\left|+\frac{1}{2}\right\rangle + c_2\left|-\frac{1}{2}\right\rangle \tag{9}$$

A quantum circuit is made of interconnections of quantum sub-circuits (e.g., quantum gates) with the following properties [1,3]: (1) quantum circuit must be reversible (reversible circuit is the circuit in which the vector of input states can be always uniquely reconstructed from the vector of output states [1,2,3]), (2) quantum circuit must have an equal number of inputs and outputs, (3) quantum circuit doesn't allow fan-out, (4) quantum circuit is constrained to be acyclic (i.e., feedback (loop) is not allowed), and (5) the transformation performed by the quantum circuit is a unitary transformation (i.e., unitary matrix). *Therefore, while each quantum circuit is reversible, not each reversible circuit is quantum* [1,3].

3 Quantum Viterbi Circuit Design

The hardware implementation for each trellis node in the Viterbi algorithm [4] requires the following components: modulo-2 adder, arithmetic adder, subtractor and

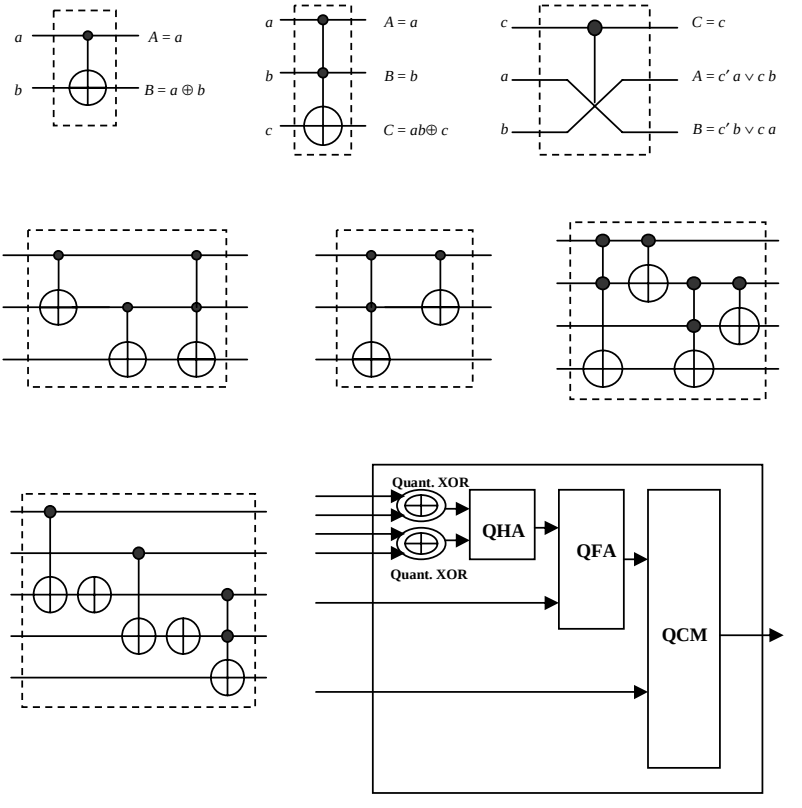


Fig. 1. Quantum circuits for the quantum realization of each trellis node in the corresponding Viterbi algorithm: (a) quantum XOR gate (Feynman gate; Controlled-NOT (C-NOT) gate), (b) quantum Toffoli gate (Controlled-Controlled-NOT (C^2-NOT) gate), (c) quantum multiplexer (Fredkin gate; Controlled-Swap (C-Swap) gate), (d) quantum subtractor (QS), (e) quantum half-adder (QHA), (f) quantum full-adder (QFA), (g) quantum equality-based comparator that compares two 2-bit numbers where an isolated XOR symbol means a quantum NOT gate, and (h) basic quantum Viterbi (QV) cell (i.e., quantum trellis node) which is made of two Feynman gates, one QHA, one QFA and one quantum comparator with multiplexing (QCM). The quantum comparator can be synthesized using a quantum subtractor and a Fredkin gate. The symbol $\oplus$ is logic XOR (exclusive OR; modulo-2 addition), $\wedge$ is logic AND, $\vee$ is logic OR, and $'$ is logic NOT.

selector (i.e., multiplexer) to be both used in one possible design of the corresponding comparator.

Figure 1 shows the various quantum circuits for the quantum realization of each quantum trellis node in the corresponding (reversible) Viterbi algorithm. Figures 1a-1c present fundamental quantum gates [1,3]. Figures 1d-1g show basic quantum arithmetic circuits of: quantum subtractor (Figure 1d), quantum half-adder (Figure 1e), quantum full-adder (Figure 1f), and quantum equality-based comparator (Figure 1g). Figure 1h introduces the basic quantum Viterbi cell (i.e., quantum trellis node) which is made of

two Feynman gates, one QHA, one QFA and one quantum comparator with multiplexing (QCM). Figure 2 shows logic circuit design of an iterative network to compare two 3-digit binary numbers: $X = x_1\ x_2\ x_3$ and $Y = y_1\ y_2\ y_3$, and Figure 3 presents the detailed synthesis of a comparator circuit which is made of a comparator cell (Figure 3a) and a comparator output circuit (Figure 3b). The extension of the circuit in Figure 2 to compare two *n*-digit binary numbers is straightforward by utilizing *n*-cells and the same output circuit.

Figure 4 illustrates the quantum circuit synthesis for the comparator cell and output circuit (which were shown in Figure 3), and Figure 5 shows the design of a quantum comparator with multiplexing (QCM) where Figure 5a shows an iterative quantum network to compare two 3-digit binary numbers and Figure 5c shows the complete design of the QCM. The extension of the quantum circuit in Figure 5a to compare two *n*-digit binary numbers is straightforward by utilizing *n* quantum cells (from Figure 4a) and the same output quantum circuit (in Figure 4b).

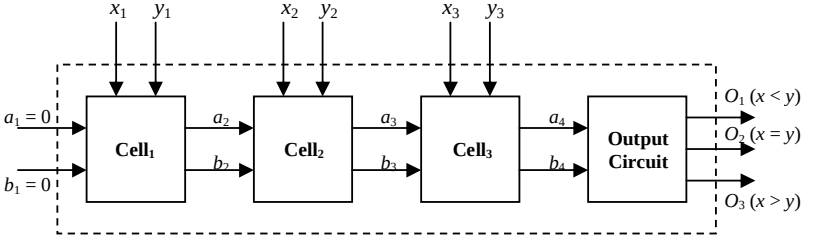


Fig. 2. An iterative network to compare two 3-digit binary numbers X and Y

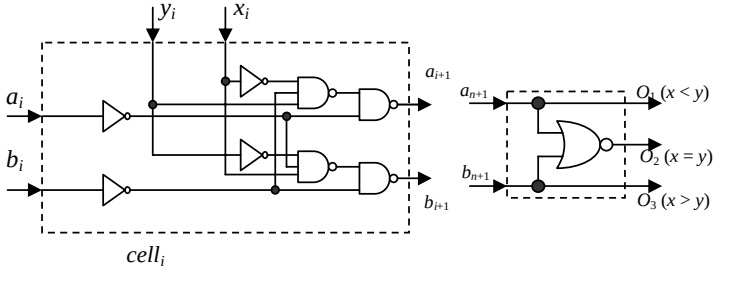


Fig. 3. Designing a comparator circuit: (a) comparator cell and (b) comparator output circuit

Figure 6 shows the complete design of a quantum trellis node (i.e., quantum Viterbi cell) that was shown in Figure 1h. The design of the quantum trellis node shown in Figure 6f proceeds as follows: (1) two quantum circuits for the first and second lines entering the trellis node each is made of two Feynman gates (i.e., two quantum XORs) to produce the difference between incoming received bits and trellis bits followed by quantum half-adder (QHA) to produce the corresponding sum (which is the Hamming distance) are shown in Figures 6a and 6b, (2) logic circuit composed of a QHA and a quantum full-adder (QFA) that adds the current Hamming distance to the previous Hamming distance is shown in Figure 6c, (3) two quantum circuits for

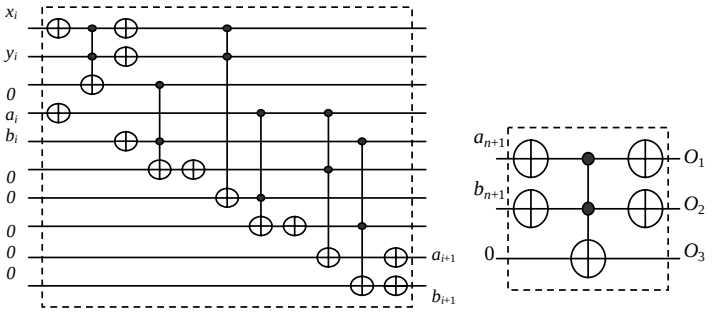


Fig. 4. Quantum circuit synthesis for the comparator cell and output circuit in Figure 3: (a) quantum comparator cell and (b) quantum comparator output circuit

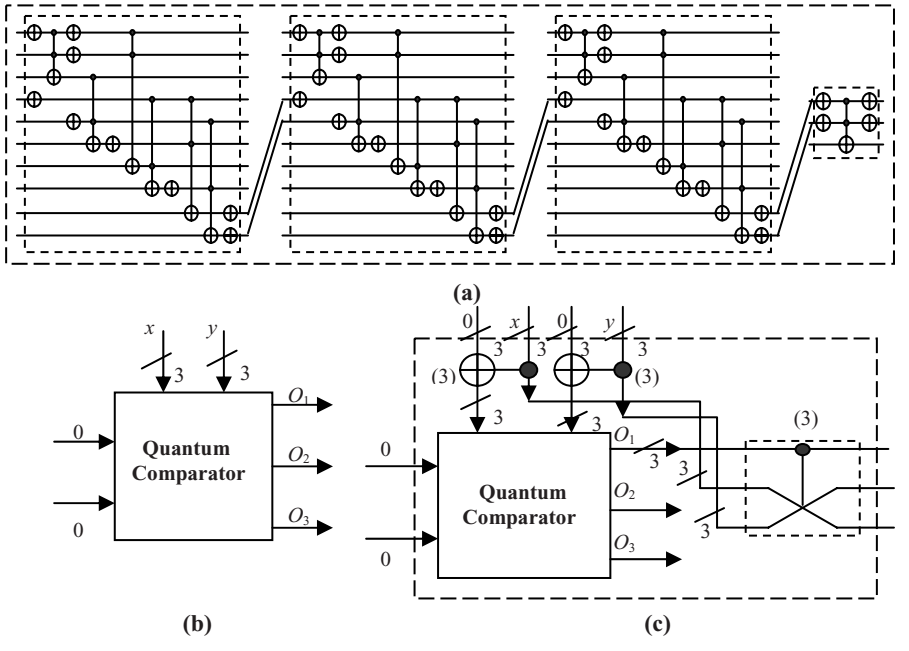


Fig. 5. Designing a quantum comparator with multiplexing (QCM): (a) an iterative quantum network to compare two 3-digit binary numbers, (b) symbol of the quantum comparator circuit in (a), and (c) complete design of QCM where the number 3 on lines indicates triple lines and (3) beside sub-circuits indicates triple circuits (i.e., three copies of each sub-circuit for the processing of the triple-input triple-output lines.)

the first and second lines entering the trellis node each is synthesized according to the logic circuit in Figure 6c (which is made of a QHA followed by a QFA) are shown in Figures 6d and 6e, (4) quantum comparator with multiplexing (QCM) in the trellis node that compares the two entering metric numbers (i.e., two entering Hamming

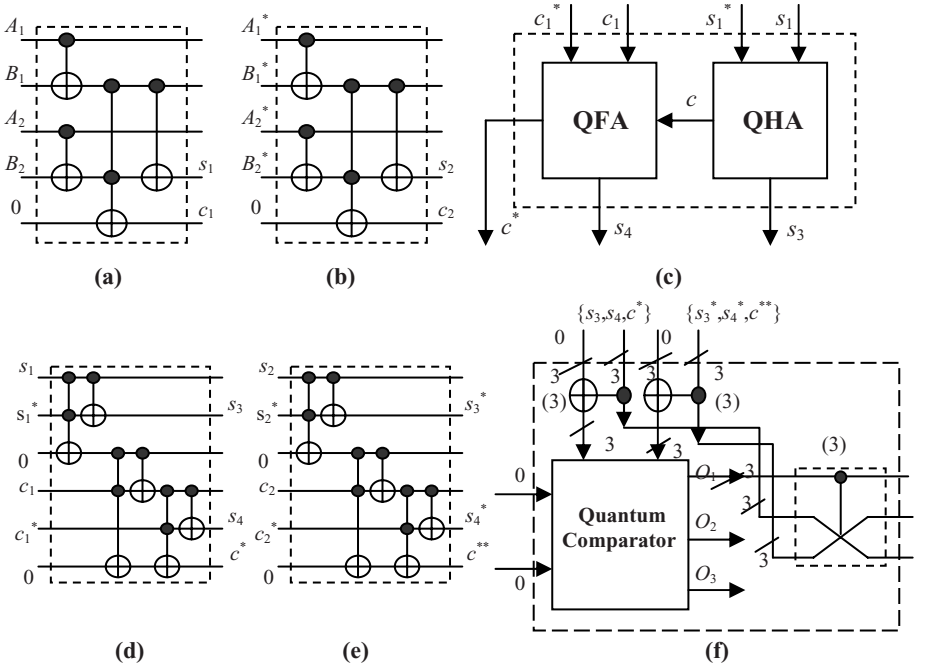


Fig. 6. The complete design of a quantum trellis node in the Viterbi algorithm that was shown in Figure 1h: (a) quantum circuit that is made of two Feynman gates to produce the difference between incoming received bits ($A_1 A_2$) and trellis bits ($B_1 B_2$) followed by quantum half-adder (QHA) to produce the corresponding sum ($s_1 c_1$) which is the Hamming distance for the first line entering the trellis node, (b) quantum circuit that is made of two Feynman gates to produce the difference between incoming received bits ($A_1^* A_2^*$) and trellis bits ($B_1^* B_2^*$) followed by quantum half-adder (QHA) to produce the corresponding sum ($s_2 c_2$) which is the Hamming distance for the second line entering the trellis node, (c) logic circuit composed of QHA and quantum full-adder (QFA) that adds the current Hamming distance to the previous Hamming distance, (d) quantum circuit in the first line entering the trellis node for the logic circuit in (c) that is made of a QHA followed by a QFA, (e) quantum circuit in the second line entering the trellis node for the logic circuit in (c) that is made of a QHA followed by a QFA, and (f) quantum comparator with multiplexing (QCM) in the trellis node that compares the two entering metric numbers: $X = s_3\, s_4\, c^*$ and $Y = s_3^*\, s_4^*\, c^{**}$ and selects using control line O_1 the path that produces the minimum entering metric (i.e., $X < Y$).

distances) and selects using the control line O_1 the path that produces the minimum entering metric (i.e., minimum entering Hamming distance) is shown in Figure 6f.

In Figures 6c-6e, the current Hamming metric $\{s_1, c_1\}$ for the first entering path of the trellis node and the current Hamming metric $\{s_2, c_2\}$ for the second entering path of the trellis node is always made of two bits (00, 01, or 10). If more than two digits (two bits) is needed to represent the previous Hamming metric for the first or second entering paths of the trellis node, then extra QFAs are added in the circuit in Figure 6c and consequently in the quantum circuits shown in Figures 6d-6e. Also, in the case when the paths entering a quantum trellis node (state) are compared and their metrics are found to be identical then a choice is made by making a guess to choose any of the

two entering paths, and this is automatically performed in the quantum circuit in Figure 6f; if ($\{s_3, s_4, c^*\} < \{s_3^*, s_4^*, c^{**}\}$) then O_1 = "1" and thus chooses $X = \{s_3, s_4, c^*\}$, else O_1 = "0" and then it chooses $Y = \{s_3^*, s_4^*, c^{**}\}$ in both cases of ($\{s_3, s_4, c^*\} > \{s_3^*, s_4^*, c^{**}\}$) or ($\{s_3, s_4, c^*\} = \{s_3^*, s_4^*, c^{**}\}$).

4 Conclusions

This paper introduces the complete synthesis of quantum circuits in the quantum domain (Q-domain) for the quantum implementation of the new quantum trellis node (i.e., quantum Viterbi cell). Quantum Viterbi design is important for the following main reasons: (1) since power reduction has become the current main concern for digital logic designers after performance (speed), quantum error-control coding is highly important for low-power circuit synthesis of future technologies such as in quantum computing (QC), and (2) super-speedy operations because of the superposition and entanglement properties that emerge in the quantum computing (QC) systems.

References

1. Al-Rabadi, A.N.: Reversible Logic Synthesis. From Fundamentals to Quantum Computing. Springer, New York (2004)
2. Landauer, R.: Irreversibility and Heat Generation in the Computational Process. IBM J. Res. and Dev 5, 183–191 (1961)
3. Nielsen, M., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Prss, Cambridge (2000)
4. Viterbi, A.J., Omura, J.K.: Principles of Digital Communication and Coding. McGraw-Hill, New York (1979)

A Hybrid Quantum-Inspired Evolutionary Algorithm for Capacitated Vehicle Routing Problem

Jing-Ling Zhang[1], Yan-Wei Zhao[1], Dian-Jun Peng[1], and Wan-Liang Wang[2]

[1] Key Laboratory of Mechanical Manufacture and Automation of Ministry of Education, Zhejiang Univ. of Tech, Hangzhou, 310032, China
[2] College of Information Engineering, Zhejiang Univ. of Tech, Hangzhou, 310032, China
zyw@zjut.edu.cn, wwl@zjut.edu.cn

Abstract. A hybrid Quantum-Inspired Evolutionary Algorithm (HQEA) with 2-OPT sub-routes optimization for capacitated vehicle routing problem (CVRP) is proposed. In the HQEA, 2-OPT algorithm is used to optimize sub-routes for convergence acceleration. Moreover, an encoding method of converting Q-bit representation to integer representation is designed. And genetic operators of quantum crossover and quantum variation are applied to enhance exploration. The proposed HQEA is tested based on classical benchmark problems of CVRP. Simulation results and comparisons with genetic algorithm show that the proposed HQEA has much better exploration quality and it is an effective method for CVRP.

Keywords: Quantum-Inspired Evolutionary Algorithm (QEA), 2-OPT, Hybrid Quantum-Inspired Evolutionary Algorithm (HQEA), Q-bit representation, Q-gate, capacitated vehicle routing problem.

1 Introduction

Since Vehicle routing problem (VRP) was firstly proposed by Dantzig and Ramser in 1959[1], it keeps focus researched in the field of operational research and combinatorial optimization. VRP is a typical NP-complete problem and can be regarded as a mixed problem of the traveling salesperson problem (TSP) and bin-packing problem (BPP). The solution of this problem is for the vehicles distribute and clients' permutation in the vehicles routes. According to different constraints VRP have many models such as Capacitated Vehicle Routing Problem (CVRP), Vehicle Routing Problem with Time Windows (VRPTW) etc. CVRP is the basic VRP, which constrains only in vehicle capacities and the maximum distance that each vehicle can travel. So far, many approaches have been proposed for CVRP. However, exact techniques are applied only to small-scale problem and the qualities of constructive heuristics are often not satisfactory. So intelligent methods have gained wide research such as genetic algorithm (GA) [2] [3], particle swarm optimization (PSO) [4] etc, and achieved better results.

QEA is based on the concept and principles of quantum computing such as a quantum bit and superposition of states. Like the evolutionary algorithms (EAs), QEA is also characterized by the representation of the individual, the evaluation function, and

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 31 – 38, 2008.
© Springer-Verlag Berlin Heidelberg 2008

the population dynamics. However, instead of binary, numeric representation, QEA uses a Q-bit as a probabilistic representation and a Q-bits individual is defined by a string of Q-bits. The Q-bit individual has the advantage that it can represent a linear superposition of states in search space probabilistically. So the Q-bit representation has a better characteristic of population diversity. The representational researches in the field of combinatorial optimization of QEA include: Narayanan [5] is the first to propose the frame of quantum-inspired genetic algorithm, in which the basic terminology of quantum mechanics is introduced in GA. And a small-scale TSP is solved successfully. Han proposed a genetic quantum algorithm [6], a quantum-inspired evolutionary algorithm [7] and a quantum-inspired evolutionary algorithm with two-phase scheme [8] for knapsack problem. Ling Wang proposed a hybrid quantum-inspired genetic algorithm for flow shop scheduling [9].

A quantum-inspired evolutionary algorithm for VRP is introduced in this paper. Firstly, the basis of QEA is analyzed in section 2. Secondly, a hybrid quantum-inspired evolutionary algorithm for CVRP is proposed in section 3. At last, a small-scale CVRP and classical big-scale benchmark problems are computed by this algorithm. Simulation results and comparisons prove that this algorithm is an effective method for VRP.

2 QEA

2.1 Representation and Strategy of Updating

In QEA, a Q-bit chromosome representation is adopted. The smallest unit of information stored in two-state quantum computer is called a Q-bit, which may be in the "1" state, in the "0" state, or in any superposition of the two. The state of a Q-bit can be represented as following:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{1}$$

Where α and β are complex numbers. $|\alpha|^2$ and $|\beta|^2$ donate the probabilities that the Q-bit will be found in the "0" state and in the "1" state respectively. Normalization of the state to unity guarantee $|\alpha|^2 + |\beta|^2 = 1$.

So a Q-bit individual with a string of m Q-bits can be expressed as following:

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ \beta_1 & \beta_2 & & \beta_m \end{bmatrix} \tag{2}$$

Where $|\alpha_i|^2 + |\beta_i|^2 = 1,\ i = 1, 2, \cdots, m$. By consequence, m Q-bits individual contains the information of 2^m states. Evolutionary computing with Q-bit representation has a better characteristic of population diversity. So QEA with small populations can obtain better results than GA with bigger populations.

In QEA, a Q-gate is an evolution operator to drive the individuals toward better solutions and eventually toward a single state. A rotation gate is often used, which is employed to update a Q-bit individual as following:

$$U(\theta_i)\begin{bmatrix}\alpha_i\\ \beta_i\end{bmatrix}=\begin{bmatrix}\cos(\theta_i) & -\sin(\theta_i)\\ \sin(\theta_i) & \cos(\theta_i)\end{bmatrix}\begin{bmatrix}\alpha_i\\ \beta_i\end{bmatrix} \tag{3}$$

Where $\left[\alpha_i,\beta_i\right]^T$ is the i-th Q-bit and θ_i is the rotation angle of each Q-bit towards either 0 or 1 state depending on its sign.

2.2 Procedure of QEA

The procedure of QEA is described in the following.

Sept 1: make $t=0$. Randomly generate an initial population $P(t)=\left\{p_1^t,\cdots,p_n^t\right\}$. Where p_j^t is a Q-bit individual, which is the j-th individual in the t-th generation. $p_j^t=\left[\begin{array}{c|c|c|c}\alpha_1^t & \alpha_2^t & \cdots & \alpha_m^t\\ \beta_1^t & \beta_2^t & & \beta_m^t\end{array}\right]$. Where m is the number of Q-bits.

Sept 2: make binary population $R(t)$ by observing the states of $P(t)$. Where $\mathrm{R}(t)=\left\{r_1^t,\cdots,r_n^t\right\}$ at generation t. One binary solution, is a binary string of length m, which is formed by selecting either 1or 0 for each bit using the probability of Q-bit, either $\left|\alpha_i^t\right|^2$ or $\left|\beta_i^t\right|^2$ of p_j^t respectively. That is, in the first place, generate a random number s between $[0,1]$, if $\left|\alpha_i^t\right|^2>s$, then set a bit in r_j^t as 1, otherwise set it as 0.

Sept 3: evaluate each solution of $R(t)$, and then record the best solution b. If the stopping condition is satisfied, output the best solution b, otherwise go on following steps.

Sept 4: apply rotation gate $U(\theta)$ to update $P(t)$.

Sept 5: set $t=t+1$, and go back to step 2.

3 HQEA for CVRP

The CVRP can be expressed by a mixed-integer programming model formulation as follows: Give $K(k=1,2\cdots K)$ vehicles at the most deliver goods for $L(i=1,2\cdots L)$ clients. If $i=0$, i is the storehouse. The carrying capacity of each vehicle is $b_k(k=1,2,\cdots K)$. The requirement of each client is $d_i(i=1,2,\cdots L)$. The transport cost from client i to client j is c_{ij}, which can be distance or fare. The optimization aim is the shortest distance which can be expressed as follows:

$$\min Z=\sum_{k=1}^{K}\sum_{i=0}^{L}\sum_{j=0}^{L} c_{ij}x_{ijk} \qquad x_{ijk}=\begin{cases}1 & \text{vehicle } k \text{ from client } i \text{ to } j\\ 0 & \text{other}\end{cases} \tag{4}$$

2-OPT sub-routes optimization algorithm is combined with QEA to develop a hybrid QEA, in which quantum crossover and quantum variation are applied to avoid getting in local extremum. The flow chart of HQEA for CVRP is illustrated in fig.1.

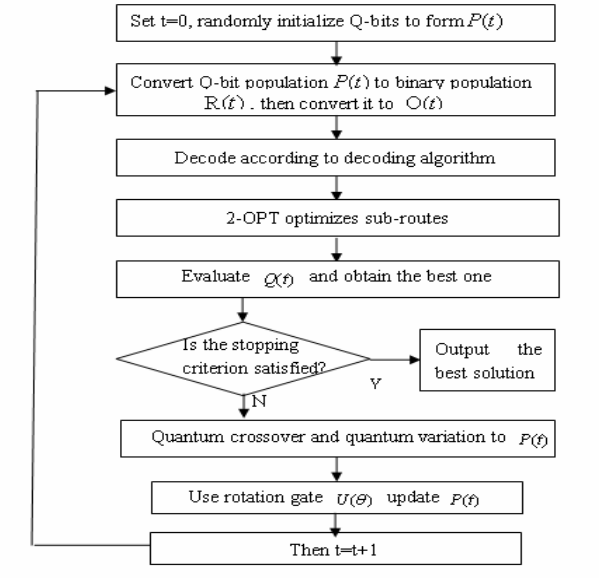


Fig. 1. Flow chart of Hybrid QEA for CVRP

The procedure of Hybrid QEA for CVRP is as follows:

1. Encoding Method

QEA of Q-bit representation only performs exploration in 0-1 hyperspace, while the solution of VRP is the permutation of all clients. Thus, the QEA cannot directly apply to VRP and it should convert Q-bit representation to permutation for evaluation. To resolve this problem, an encoding method is designed to convert Q-bit representation to integer representation, which can be summarized as follows: at the first place, convert Q-bit representation to binary representation by observing the value of α_i or β_i. And then the binary representation stands for random key representation. Finally, integer representation can be constructed based on the random key representation.

The encoding procedure is as follows:

Step 1: use the encoding method in paper [10]. To serve a problem of L clients and K vehicles, use an integer vector $(g_1, g_2, \cdots g_{L*K})$ from 1 to $L*K$ to express the individual r_j. Where $g_j (j=1,2\cdots L*K)$ is the j-th gene in the integer vector. The relation of client i and vehicle k can be confirmed in formulations (5) and (6). That is, by

evaluating the two formulations client i may be served by vehicle k, whether or not it should be checked up by restriction.

$$i = \left(g_j - \left[\frac{g_j - 1}{L} \right] \times L \right) \qquad i \in (1, 2 \cdots L) \tag{5}$$

$$k = \left(\left[\frac{g_j - 1}{L} \right] + 1 \right) \qquad k \in (1, 2, \cdots K) \tag{6}$$

Where [ ] stand for the operator of getting integer.

Step 2: convert the individual r_j expressed by the integer vector $(g_1, g_2, \cdots g_{L*K})$ to Q-bit individual, which can be summarized as follows: each gene $g_j (j = 1, 2 \cdots L*K)$ in the integer vector $(g_1, g_2, \cdots g_{L*K})$ is expressed as Q-bits string of length n, where $2^n \geq L*K$. So the Q-bit individual of length $L*K*n$ is obtained.

2. Decoding Procedure

Convert Q-bit individual of length $L*K*n$ to binary individual of length $L*K*n$ by the method in part 2.2. And then convert the binary individual to an integer vector $(g_1, g_2, \cdots g_{L*K})$ of different integer from 1 to $L*K$. To keep the values different in $(g_1, g_2, \cdots g_{L*K})$, a method as follows is used: if two values in $(g_1, g_2, \cdots g_{L*K})$ are different, let the bigger value denote the bigger number. Otherwise, let the one first appears denote the smaller number. The clients' permutation in the vehicles routes can be confirmed by a decoding algorithm in the following:

Step1. Set $x_{ijk} = 0, y_{ik} = 0$。 $j = 1$, $b'_k = b_k$, $n_k = 1$, $R_k = (0)$.

Step2. Evaluate the value of i and k by formulations (5) (6).

Step3. If $y_{ik} = 0$, indicate client i is not served, and then judge $d_i \leq b'_k$, if satisfy, then $y_{ik} = 1$, $b'_k = b'_k - d_i$, $x_{iR_k^{n_k}k} = 1$, $n_k = n_k + 1$. Otherwise, perform Step4.

Step4. Set $j = j + 1$, go back to Step2 and perform the loop to $j = L*K$。And then if all $y_{ik} = 1$, indicate the result is a viable solution, otherwise it is not a viable solution.

Where b'_k is the carrying capacity of vehicle k, n_k is the number of client served by vehicle k, R_k is the set of client in, $R_k^{n_k}$ is the n_k -th client in vehicle k.

3. Initialize Population

Randomly generate an initial population. That is, randomly generated any value in $[0,1]$ for α_i and β_i.

4. 2-OPT Sub-routes Optimization Algorithm

2-OPT is a heuristic algorithm for TSP in common use, by which the permutation of clients in each sub-route can be optimized locally to obtain the best solution.

5. Fitness Evaluation
The fitness of each solution can be related to the objective function value. Use formulation (4) to gain the objective function value Z. If the solution is impossible, set a big integer to Z. And then make the fitness function $Fitness = 1/Z$.

6. Quantum Crossover
The probability of crossover is P_c. Use one point crossover as follows: randomly determinates a position i, and then the Q-bits of the parents before position i are reserved while the Q-bits after position i are exchanged.

7. Quantum Variation
Select an individual with P_m. Randomly selects a position i in this individual, and then exchange α_i and β_i.

4 Experiment Results and Comparisons

4.1 Testing Problem 1

In this paper, a CVRP of 8 clients, 2 vehicles and 1 storehouse is chosen [10]. The carrying capacity of each vehicle is 8 ton. The optimization result is 67.5 by this algorithm, and the clients' permutations in the vehicle routes are (0 - 2 - 8 - 5 - 3 - 1 – 0) and (0 - 6 - 7 - 4 – 0), which is the same to the known optimization result.

In this algorithm, set population size as 10, maximum generation as 100, crossover probability as 1, and mutation probability as 0.03. The simulation result is showed in Fig.2. From the figure, we can see the proposed HQEA can obtain the optimization result with short time.

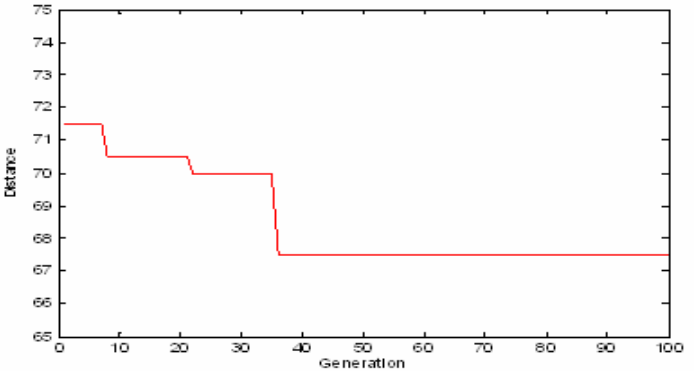


Fig. 2. Simulation result of HQEA for CVRP

Then, run the algorithm 20 times for the above problem, and compare with double population genetic algorithm in paper [3], whose results is superior to standard genetic algorithm. The 20 times results of comparison are listed in Table 2. And the

Table 2. Comparisons of HQEA and Double populations GA

Experiment times	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
HQEA	67.5	70	67.5	70	69	70	67.5	69	69	70
Double populations GA	70	69.5	67.5	71	69	70.5	72	67.5	71.5	69
Experiment times	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
HQEA	70.5	67.5	67.5	70	67.5	69	69	67.5	67.5	67.5
Double populations GA	67.5	69	71	70	67.5	70.5	69	69.5	71	69

Table 3. The Statistic of simulation results

	Average value	Times of search optimization	Probability of gain optimization (%)
HQEA	68.4375	9	45%
Double populations GA	69.575	4	20%

statistical results are summarized in Table 3. From Table 3 we can see the probability of gaining the optimization by HQEA is better than that by double populations GA.

4.2 Testing Benchmark

To test the applicability of HQEA for big-scale CVRP, several typical benchmark problems in CVRP storage are chosen, and compare the performance of genetic algorithm in paper [11]. The comparison results are summarized in Table 4. In HQEA, set population size as 20, maximum generation as 1000, crossover probability as 1, and mutation probability as 0.03. The average values are the results of randomly performing 10 times. From Table 4, it can be seen that the search results of HQEA are better than that of HGA.

Table 4. Comparisons of HQEA and HGA

Benchmark (the best value)	**HQEA**			**HGA**		
	best	average	worst	best	average	worst
E-n22-k4 (375)	375	376.125	382	375	376.625	382
E-n33-k4 (835)	847	851.5	895	867	894.3	925

5 Conclusions

This paper proposes a hybrid QEA with 2-OPT sub-routes optimization algorithm for CVRP. In this algorithm, genetic operators of quantum crossover and quantum variation are introduced to enhance exploration. 2-OPT algorithm for optimizing sub-routes can enhance the convergence speed. The encoding method of converting

integer representation to Q-bit representation makes QEA applicable for scheduling problem. Simulation results and comparisons with GA show that the HQEA proposed has better convergence stability and calculate efficiency, and so the HQEA is an effective method for CVRP.

QEA can be applied in other model such as: open vehicle routing problem and dynamic vehicle routing problem, which are the future research work.

Acknowledgements. This paper is supported by the National Natural Science Foundation of China (Grant No.60573123), The National 863 Program of China (Grant No.2007AA04Z155), and the Plan of Zhejiang Province Science and Technology (Grant No. 2007C21013).

References

1. Dantzig, G., Ramser, J.: The Truck Dispatching Problem[J]. Management Science (6), 80–91 (1959)
2. Zhang, L.P., Chai, Y.T.: Improved Genetic Algorithm for Vehicle Routing Problem[J]. Systems Engineering- Theory & Practices 22(8), 79–84 (2002)
3. Zhao, Y.W., Wu, B.: Double Populations Genetic Algorithm for Vehicle Routing Problem [J]. Computer Integrated Manufacturing Systems 10(3), 303–306 (2004)
4. Wu, B., Zhao, Y.W., Wang, W.L., et al.: Particle Swarm Optimization for Vehicle Routing Problem [C]. In: The 5th World Congress on Intelligent Control and Automation, pp. 2219–2221 (2004)
5. Narayanan, A., Moore, M.: Quantum-Inspired Genetic Algorithms. In: Proc. of the 1996 IEEE Intl. Conf. on Evolutionary Computation (ICEC 1996), Nogaya, Japan, IEEE Press, Los Alamitos (1996)
6. Han, K.H.: Genetic Quantum Algorithm and its Application to Combinatorial Optimization Problem. In: IEEE Proc. of the 2000 Congress on Evolutionary Computation, San Diego, USA. IEEE Press, Los Alamitos (2000)
7. Han, K.H., Kim, J.H.: Quantum-inspired Evolutionary Algorithm for a Class of Combinatorial Optimization [J]. IEEE Trans on Evolutionary Computation (2002)
8. Han, K.H., Kim, J.H.: Quantum-inspired Evolutionary Algorithms with a New Termination Criterion H, Gate, and Two-Phase Scheme [J]. IEEE Trans on Evolutionary Computation (2004)
9. Wang, L., Wu, H.: A Hybrid Quantum-inspired Genetic Algorithm for Flow Shop Scheduling. In: Huang, D.-S., Zhang, X.-P., Huang, G.-B. (eds.) ICIC 2005. LNCS, vol. 3645, pp. 636–644. Springer, Heidelberg (2005)
10. Jiang, D.L., Yang, X.L.: A Study on the Genetic Algorithm for Vehicle Routing Problem [J]. Systems Engineering-Theory & Practice 19(6), 44–45 (1999)
11. Jiang, C.H., Dai, S.G.: Hybrid Genetic Algorithm for Capacitated Vehicle Routing Problem. Computer Integrated Manufacturing Systems. 13(10), 2047–2052 (2007)

Improving Tumor Clustering Based on Gene Selection

Xiangzhen Kong[1], Chunhou Zheng [1,2,*] , Yuqiang Wu[3], and Yutian Wang[1]

[1] School of Information and Communication Technology, Qufu Normal University, Rizhao, Shandong, 276826, China
kongxzhen@163.com
[2] Intelligent Computing Lab, Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei, Anhui, China
Zhengch99@126.com
[3] Institute of Automation, Qufu Normal University

Abstract. Tumor clustering is becoming a powerful method in cancer class discovery. In this community, non-negative matrix factorization (NMF) has shown its advantages, such as the accuracy and robustness of the representation, over other conventional clustering techniques. Though NMF has shown its efficiency in tumor clustering, there is a considerable room for improvement in clustering accuracy and robustness. In this paper, gene selection and explicitly enforcing sparseness are introduced into clustering process. The independent component analysis (ICA) is employed to select a subset of genes. The unsupervised methods NMF and its extensions, sparse NMF (SNMF) and NMF with sparseness constraint (NMFSC), are then used for tumor clustering on the subset of genes selected by ICA. The experimental results demonstrate the efficiency of the proposed scheme.

Keywords: Gene Expression Data, Clustering, Independent Component Analysis, Non-negative Matrix Factorization.

1 Introduction

Till now, many techniques have been proposed and used to analyze gene expression data and they have demonstrated the potential power for tumor classification [1]-[3], [4]. Though the NMF based clustering algorithms are useful, one disadvantage of them is that they cluster the microarray dataset with thousands of genes directly, which makes the clustering result not very satisfying. One important reason is that the inclusion of irrelevant or noisy variables may also degrade the overall performances of the estimated classification rules. To overcome this problem, in this paper we propose to perform gene selection before clustering to reduce the effect of irrelevant or noisy variables, so as to achieve a better clustering result. Gene selection has been used for cell classification [5], [6]. Gene selection in this context has the following biological explanation: most of the abnormalities in cell behavior are due to irregular gene activities and thus it is critical to highlight these particular genes.

In this paper, we first employ ICA to select a subset of genes, and then apply the unsupervised method, NMF and its extensions, SNMF and NMFSC, to cluster the

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 39 – 46, 2008.
© Springer-Verlag Berlin Heidelberg 2008

tumors on the subset of genes selected by ICA. To find out which cluster algorithm cooperates best with the proposed gene selection method, we compared the results using NMF, SNMF and NMFSC, respectively, on different gene expression datasets.

2 Gene Selection by ICA

2.1 Independent Component Analysis (ICA)

ICA is a useful extension to principal component analysis (PCA) and it was originally developed for blind separation of independent sources from their linear mixtures [7]. Unlike PCA that is to decorrelate the dataset, ICA aims to make the transformed coefficients be mutually independent (or as independent as possible).

Considering an $p \times n$ data matrix X, whose rows r_i ($i = 1, \cdots, p$) represent the individuals (genes for expression data) and whose columns c_j $(j = 1, \cdots, n)$ are the variables (cell samples for expression data), the ICA model of X can be written as:

$$X = SA \tag{1}$$

Where A is a $n \times n$ mixing matrix. s_k, the columns of S, are assumed to be statistically independent and named the ICs of S. Model (1) implies that the columns of X are linear mixtures of the ICs. The statistical independence between s_k can be measured by mutual information $I = \sum_k H(s_k) - H(S)$, where $H(s_k)$ is the marginal entropy of the variable s_k and $H(S)$ is the joint entropy of S. Estimating the ICs can be accomplished by finding the right linear combinations of the observational variables. We can invert the mixing matrix such that

$$S = XA^{-1} = XW \tag{2}$$

Then ICA algorithm is used to find a projection matrix W such that the rows of S are as statistically independent as possible. Several algorithms have been proposed to implement ICA [8], [9]. In this paper, we employ the FastICA algorithm [9] to model the gene expression data.

2.2 Gene Selection: An ICA Based Solution

The selection is performed by projecting the genes onto the desired directions obtained by ICA. In particular, the distribution of gene expression levels on a cell is "approximately sparse", with heavy tails and a pronounced peak in the middle. Due to this, the projections obtained by ICA should emphasize this sparseness. Highly induced or repressed genes, which may be useful in cell classification, should lie on the tails of the distributions of s_j $(j = 1, \cdots, z)$, where z is the actual number of ICs we estimated from gene expression data. Since these directions are independent, they may catch different aspects of the data structure that could be useful for classification tasks [5].

The proposed gene selection method is based on a ranking of the p genes. This ranking is obtained as follows [5]:

Step1. z independent components $s_1, \cdots, s_z$ with zero mean and unit variance are extracted from the gene expression data set using ICA;

Step2. For gene l ($l = 1, \cdots, p$), the absolute score on each component $| s_{lj} |$ is computed. These z scores are synthesized by retaining the maximum one, denoted by $g_l = \max_j | s_{lj} |$;

Step 3. All of the p genes are sorted in increasing order according to the maximum absolute scores $\{g_1, \cdots g_p\}$ and for each gene the rank $r(l)$ is computed.

In our experiments, we also found that ICA is not always reproducible when used to analyze gene expression data. In addition, the results by an ICA algorithm are not "ordered" because the ICA algorithm may converge to local optima [10]. To solve this problem, we run the independent source estimation 100 times with different random initializations. In each time, we choose the subset of the last m genes (with $m << p$). The rationale behind this is that these m genes show, with respect to at least one of the components, a behavior across the cells that differs most from that of the bulk of the genes [5]. After running the algorithm 100 times, we obtain $100 \times m$ genes, most of them are reduplicate. The selected genes are then sorted in descending order according to the selected frequency of each gene. Finally, we selected the first m genes for the next step of clustering.

3 Clustering with NMF

The NMF methods resort to factor the gene-expression matrix Y into the product of two matrices of non-negative entries: $Y \approx VH$. In NMF model, each entry v_{ij} in V is the coefficient of gene i in metagene j and each entry h_{ij} in H represents the expression level of metagene i in sample j. In such a factorization, matrix H can be used to group the n samples into k clusters [2]. In literature [2], each cell sample is placed into a cluster corresponding to the most highly expressed metagene in the sample, i.e. sample j is placed in cluster i if h_{ij} is the largest entry in column j.

NMF can reduce the dimensionality of expression data from thousands of gene to metagenes. However, standard NMF can not control the sparseness of the decomposition and thus does not always yield a parts-based representation. solve this problem, Hoyer [11] extended the NMF framework to include an adjustable sparseness parameter. SNMF and NMFSC are extensions of those ideas [11], [12]. The main improvement of them is that the sparseness can be adjusted explicitly, rather than implicitly. It has been shown that the extension of NMF, e.g. SNMF and NMFSC, coupled with the model selection mechanism [2], could improve cancer class discovery on the microarray datasets [3], [4]. The detail algorithm of SNMF and NMFSC can be found in [4] and [14].

4 Experimental Results

We apply the proposed method to two public available datasets: Leukemia dataset [1] and central nervous system embryonal tumors [13]. We first employ ICA to select m genes, and then apply the NMF and its extensions, SNMF and NMFSC, to clustering the tumors using the selected genes. We run the gene selection procedure with z ranging from 1 to 10. For each value of z, we tried 40 different values of m ranging from 1 to p. To demonstrate the efficiency of the proposed strategy, we applied NMF, SNMF and NMFSC respectively, to the subsets of selected genes for cancer class discovery.

4.1 Leukemia Dataset

The Leukemia dataset has become a benchmark in the cancer classification community. In this dataset, the distinction between acute myelogenous leukemia (AML) and acute lymphoblastic leukemia (ALL), as well as the division of ALL into T and B cell subtypes, is well known. The dataset contains $p = 5{,}000$ genes in 38 cells and consists of 19 cases of B_cell acute lymphoblastic leukemia (ALL_B), 8 cases of T_cell acute lymphoblastic leukemia (ALL_T) and 11 cases of AML.

Fig. 1 and Fig. 2 display the results obtained with $z = 6$ for the number of ICs and 40 different values of m for the number of selected genes. The clustering accuracy rate for each data set is listed in Table 1 and Table 2. For comparison, we also list the results of [2] and [4] in the two figures and tables, i.e. the result of NMF, SNMF, and NMFSC with $m = 5{,}000$. From the two Figures we can see that no matter the rank is 2 or 3, gene selection by suitable projections can achieve commensurate or better performance than the methods without gene selection. At rank $k = 2$, SNMF achieves

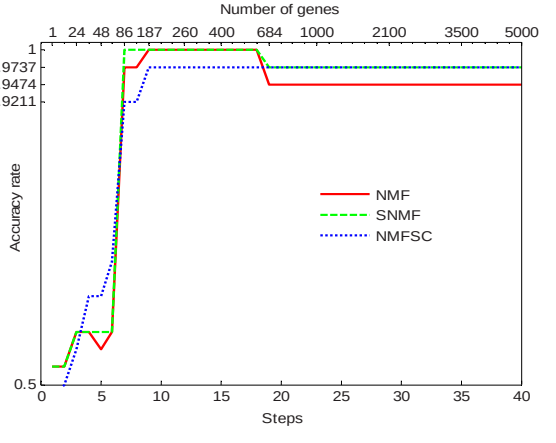


Fig. 1. Leukemia data set: clustering accuracy rate. The axis at the top of the plot indicates the number of genes retained at each step for the number of m ($z = 6$ for ICA, $k = 2$ for NMF and its extensions).

Table 1. Leukemia Data Set: Accuracy rates for different values of m ($z = 6$ for ICA, $k = 2$ for NMF and its extensions)

m	1	24	86	187	400	550	684	1000	2100	5000
NMF	0.5000	0.5789	0.9737	1.0000	1.0000	1.0000	0.9474	0.9474	0.9474	0.9474
SNMF	0.5000	0.5789	1.0000	1.0000	1.0000	0.9737	0.9737	0.9737	0.9737	0.9737
NMFSC	0.4737	0.5526	0.9211	0.9737	0.9737	0.9737	0.9737	0.9737	0.9737	0.9737

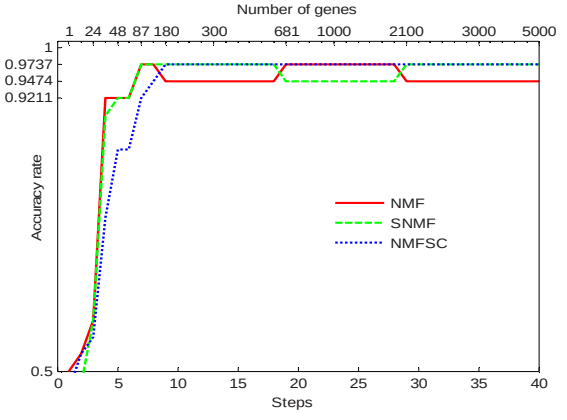


Fig. 2. Leukemia data set: clustering accuracy rate. The axis at the top of the plot indicates the number of genes retained at each step. ($z = 6$ for ICA, $k = 3$ for NMF and its extensions).

Table 2. Leukemia Data Set: Accuracy rates for different values of m ($z = 6$ for ICA, $k = 3$ for NMF and its extensions)

m	1	48	87	100	187	400	681	1000	2100	5000
NMF	0.5000	0.9211	0.9737	0.9737	0.9474	0.9474	0.9737	0.9737	0.9474	0.9474
SNMF	0.4737	0.9211	0.9737	0.9737	0.9737	0.9737	0.9737	0.9737	0.9737	0.9737
NMFSC	0.4737	0.8421	0.9474	0.9474	0.9737	0.9737	0.9737	0.9737	0.9474	0.9737

100% accuracy rate when the number of retained genes is 86; for NMF, it can also achieve 100% accuracy rate with the number of key genes being 183; for NMFSC, it could achieve the accuracy rate by 97.37% with 198 genes (misclassified AML_13 to ALL). At rank $k = 3$, although the accuracy rates are not improved compared with those in [2], [4], the proposed algorithm could achieve 97.37% accuracy rate (AML_13 sample was misclassified to ALL_B) with less key genes. The smallest number of genes which yields the largest accuracy rate is listed in Table 3. On this dataset, the ICA-based genes selection makes the class structure more evident, i.e., the use of suitable subsets of genes instead of the whole set could yield better classification performance.

Table 3. Performance comparisons of NMF and its extensions. In this table, we list the smallest number of genes which yields the largest accuracy rate ($z = 6$ for ICA). At rank $k = 2$, NMF and SNMF split the samples into two subtypes with no mistake. NMFSC ($S_v = 0.1, S_h = 0$) misclassified AML_13 to ALL. However, at rank $k = 3$, they all misclassified the same AML (AML_13) sample to ALL_B.

k	Algorithm	The smallest number of genes	Accuracy rate
2	NMF	183	100%
	SNMF	86	100%
	NMFSC	198	97.37%
3	NMF	87	97.37%
	SNMF	87	97.37%
	NMFSC	187	97.37%

4.2 Central Nervous System Tumors

This dataset is composed of four types of central nervous system embryonal tumors [13]. The dataset used in our experiment contains $p = 5{,}560$ genes in 34 samples representing four distinct morphologies: 10 classic medulloblasomas, 10 malignant gliomas, 10 rhabdoids and 4 normals. In previous studies [2] with no gene selection, NMF and SNMF, suggested a four-cluster split with high cophenetic coefficient, and NMF method made two mistakes. One mistake is to assign a glioma (Brain_MGlio_8) to rhabdoid. However, the other mistake is more serious. It incorrectly assigned a rhabdoid sample (Brain_Rhab-10) to normal. Such an assignment will definitely delay the treatment, if at all, of the patient and thus highly undesired. The SNMF method correctly split the samples into four clusters stably with high cophenetic coefficient [4]. It made only one mistake: the same glioma (Brain_MGlio_8) was assigned to rhabdoid. Significantly, It correctly clusters the four normal samples into a distinct group.

In our study, we also adopted the four-cluster suggestion as in [4], i.e. rank=4 for the NMF and its extensions. Through the gene selection strategy, we obtained the clustering results which made only one mistake with high cophenetic coefficient when applying the NMF and its extensions on the appropriate selected key genes. The mistake was the same as that in [4], i.e. the glioma (Brian_MGlio_8) was assigned to rhabdoid. We also compared the clustering accuracy rate for each subset of the genes in

Table 4. Central Nervous System Tumors Data Set: Accuracy rates for different values of m ($z = 7$ for ICA, $k = 4$ for NMF and its extensions)

m	1	75	170	220	450	687	847	1000	3000	5560
NMF	0.3824	0.8529	0.9412	0.9412	0.9412	0.9412	0.9706	0.9706	0.9412	0.9412
SNMF	0.3824	0.8824	0.9412	0.9706	0.9706	0.9706	0.9706	0.9706	0.9706	0.9706
NMFSC	0.3824	0.8529	0.9118	0.9412	0.9412	0.9706	0.9706	0.9706	0.9412	0.9706

Table 5. Performance comparisons of NMF and its extensions. In this table, we list the smallest number of genes which yields the largest accuracy rate ($z = 7$ for ICA). At rank 4 ($k = 4$), NMF, SNMF and NMFSC all split the samples into four subtypes with only one mistake made (a Brain_MGlio_8 was assigned to rhabdoid) .

k	Algorithm	The smallest number of genes	Accuracy rate
4	NMF	847	97.06%
	SNMF	220	97.06%
	NMFSC	687	97.06%

Table 4 when using different clustering algorithm. The smallest number of genes which yields the largest accuracy rate is listed in Table 5. From the result described above, it can be seen that, for this data set, both sparseness and gene selection could improved the clustering result, yet for SNMF and NMFSC, gene selection can not give better result.

5 Conclusion and Future Work

In this paper, we employed ICA to model the gene expression data for gene selection and then applied NMF and its extensions, SNMF and NMFSC, for cancer clustering using the selected genes. To validate the proposed method, we applied it to the leukemia and brain cancer datasets. The improved clustering results were achieved by selecting the key genes with ICA. We selected the suitable values for both the number of components for ICA and the number of genes by experiment, and found out the smallest number of genes which yielded maximum accuracy rate comparatively applying different clustering algorithm.

From the experimental results we can conclude that the gene selection is useful to detect the subsets of relevant genes for tumor clustering, which was especially embodied in NMF based method. For NMFSC, it appears that gene selection can not improve the clustering result. As regards the influence of gene selection to clustering stability, it can not find obvious rules. In future, systematic studies on larger datasets need to be conducted for more convincing arguments.

Acknowledgements. This work was supported by the grants of the National Science Foundation of China, No. 30700161, China Postdoctoral Science Foundation, No. 20070410223 and Scientific Research Startup Foundation of Qufu Normal University, No. Bsqd2007036.

References

1. Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S.: Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science 286, 531–537 (1999)

2. Brunet, J.-P., Tamayo, P., Golub, T.R., Mesirov, J.P.: Metagenes and Molecular Pattern Discovery Using Matrix Factorization. Proc. Natl. Acad. Sci. USA 101(12), 4164–4416 (2004)
3. Kong, X.Z., Zheng, C.H., Wu, Y.Q., Shang, L.: Molecular Cancer Class Discovery Using Non-negative Matrix Factorization with Sparseness Constraint. In: Huang, D.-S., Heutte, L., Loog, M. (eds.) ICIC 2007. LNCS, vol. 4681, pp. 792–802. Springer, Heidelberg (2007)
4. Gao, Y., George, C.: Improving Molecular Cancer Class Discovery Through Sparse Non-negative Matrix Factorization. Bioinformatics 21, 3970–3975 (2005)
5. Daniela, G., Calò., G.G., Marilena, P., Cinzia, V.: Variable Selection in Cell Classification Problems: A Strategy Based on Independent Component Analysis. Studies in Classification. Data Analysis, and Knowledge Organization, Part I, 21–29 (2006)
6. Pan, W.: A Comparative Review of Statistical Methods for Discovering Differently Expressed Genes in Replicated Microarray Experiments. Bioinformatics 18, 546–554 (2002)
7. Hyvärinen, A., Oja, E.: Independent Component Analysis: Algorithm and Applications. Neural Netw. 2000(13), 411–430 (2000)
8. Huang, D.S., Zheng, C.H.: Independent Component Analysis-based Penalized Discriminant Method for Tumor Classification Using Gene Expression Data. Bioinformatics 22, 1855–1862 (2006)
9. Hyvärinen, A.: Fast and Robust Fixed-point Algorithms for Independent Component Analysis. IEEE Trans. Neural Netw. 10, 626–634 (1999)
10. Chiappetta, P., Roubaud, M.C., Torresani, B.: Blind Source Separation and the Analysis of Microarray Data. J. Comput. Biol. 11, 1090–1109 (2004)
11. Hoyer, P.O.: Non-negative Sparse Coding Neural Networks for Signal Processing. In: Proceedings of IEEE Workshop on Neural Networks for Signal Processing, Martigny, Switzerland, pp. 557–565 (2002)
12. Hoyer, P.O.: Non-negative Matrix Factorization with Sparseness Constraints. J. Mach. Learn. Res. 5, 1457–1469 (2004)
13. Pomeroy, S.L., Tamayo, P., Gaasenbeek, M., Sturla, L.M., Angelo, M., McLaughlin, M.E., Kim, J.Y., Goumneroval, L.C., Clack, P.M., Lau, C., Allen, J.C.: Prediction of Central Nervous System Embryonal Tumour Outcome Based on Gene Expression. Nature 415, 436–442 (2002)

A Hybrid Intelligent Algorithm for the Vehicle Routing with Time Windows

Qiuwen Zhang[1], Tong Zhen[1], Yuhua Zhu[1], Wenshuai Zhang[2], and Zhi Ma[1]

[1] College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China
[2] Henan Vocational College of Agriculture, Zhongmu 451450, China
zhangqwen@sohu.com

Abstract. Vehicle routing problems (VRP) arise in many real-life applications within transportation and logistics. This paper considers vehicle routing models with time windows and its hybrid intelligent algorithm. Vehicle routing problem with time windows (VRPTW) is an NP-complete optimization problem. The objective of VRPTW is to use a fleet of vehicles with specific capacity to serve a number of customers with fixed demand and time window constraints. A hybrid intelligent algorithm base on dynamic sweep and ant colony algorithm (DSACA-VRPTW) is proposed to solve this problem. Firstly, each ant's solution might be improved by dynamic sweep algorithm. Then a new improved ant colonies technique is proposed for it. Finally, Solomon's benchmark instances (VRPTW 100-customer) are tested for the algorithm and shows that the DSACA is able to find solutions for VRPTW.

Keywords: ant colony algorithm (ACA); dynamic sweep (DS); vehicle routing problem with time window (VRPTW); Optimization.

1 Introduction

The Vehicle Routing Problem (VRP) has been largely studied because of the importance of mobility in logistic and supply-chains management that relies on road network distribution. Many different variants of this problem have been formulated to provide a suitable application to a variety of real world cases, with the development of advanced logistic systems and optimization tools. The features that characterize the different variants aim on one hand to take into account the constraints and details of the problem, while on the other to include different aspects of its nature, like its dynamicity, time dependency and/or stochastic aspects. The richness and difficulty of this type of problem, has made the vehicle routing an area of intense investigation.

In this paper we focus on the Vehicle Routing with Time Windows, The Vehicle Routing with Time Windows Problem consists in optimally routing a fleet of vehicles of fixed capacity when travel times are time dependent, in the sense that the time employed to traverse each given arc, depends on the time of the day the travel starts from its originating node. The optimization method consists in finding solutions that minimize two hierarchical objectives: the number of tours and the total travel time. Optimization of total travel time is a continuous optimization problem that in our

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 47 – 54, 2008.
© Springer-Verlag Berlin Heidelberg 2008

approach is solved by discretizing the time space in a suitable number of subspaces. New time dependent local search procedures are also introduced, as well as conditions that guarantee that feasible moves are sought for in constant time.

2 Problem Description

The vehicle routing problem has been an important problem in the field of distribution and logistics since at least the early 1960s [1]. It is described as finding the minimum distance or cost of the combined routes of a number of vehicles m that must service a number of customers n. Mathematically, this system is described as a weighted graph G=(V, A, d) where the vertices are represented by V={ i }, i =0,1, ···,n, and the arcs are represented by $A=\{(v_i, v_j): i \neq j\}$. A central depot where each vehicle starts its route is located at v_0 and each of the other vertices represents the n customers. The distances associated with each arc are represented by the variable d_{ij} which is measured using Euclidean computations. Each customer is assigned a non-negative demand q_i and each vehicle is given a capacity constraint Q_k . The problem is solved under the following constraints.

$$\min F = \{\sum_i \sum_j \sum_k d_{ij} x_{ij}^k\}, \tag{1}$$

$$\sum_i q_i y_t^k \le Q_k, \forall k, \tag{2}$$

$$\sum_i y_t^k = 1, i \in v, \tag{3}$$

$$\sum_i x_{ij}^k = y_j^k, j \in v, \forall k \quad \sum_j x_{ij}^k = y_i^k, i \in v, \forall k, \tag{4}$$

$$\sum_{i,j \in s \times s} \sum x_{ij}^k \le |S| - 1, S \subseteq v, 2 \le |S| \le n-1 \quad x_{ij}^k \in (0,1), y_{ij}^k = (0,1), \tag{5}$$

$$e_i + t_{ij} \le s_j \quad i, j \in 1, \cdots, n, \tag{6}$$

$$\min M = \sum_{k=1}^{m} q_i / Q_k, \tag{7}$$

$$\min N = \{\sum_i \sum_j \sum_k d_{ij} x_{ij}{}^k q_i\}, \tag{8}$$

$$Z = \min(W_1 * F + W_2 * M + W_3 * N)$$
$$= \min(W_1 * \sum_{k=1}^{m} q_i \Big/ Q_i + W_2 * \sum_i \sum_j \sum_k d_{ij} x_{ij}^k + W_3 * \sum_i \sum_j \sum_k d_{ij} x_{ij}^k)' \quad (9)$$

where objective function(1) is transports the way to be shortest; (2) is the vehicles limit; (3) guaranteed that each vehicles visit one time to each fixed point; (4), (5), is for can the looping; (6) is with time windows demand. S for G fixed-point integer. Objective function (7), (8), is the number of cars that use bulk at least, the minimum distance transport. On the three objectives of this function (1), (7), (8), the problem of logistic transporting multiple-objective optimization order: First the smallest of vehicles traveling path, that is, to minimize the number of vehicles required. Then the vehicles for transporting is least optimize the size of the total distance traveled and waiting time. To solve the problem need, we will take these issues into multi-target single goals. The multi-objective problems transform a single objective problem by various means. The logistic transportation problems in this have objective function selected as (9), where, W1, W2 and W3 is the control factor, different values for different goals of our attention, the application will be in accordance with logistic transportation to determine the actual situation.

3 Ant Colony Optimization

Ant Colony Optimization (ACO) was introduced by Dorigo et al. in [2], and it is based on the idea that a large number of simple artificial agents are able to build solutions via low-level based communication, inspired by the collaborative behavior of ant colonies. A variety of ACO algorithms has been proposed for discrete optimization, as discussed in[3], and have been successfully applied to the traveling salesman problem, symmetric and asymmetric[2,7], the quadratic assignment problem[9],graph-coloring problem[6], job-shop/flow-shop[5], sequential ordering[8], vehicle routing[4].ACO can be applied to optimization problems based on graphs. The basic idea is to use a positive feedback mechanism to reinforce those arcs of the graph that belong to a good solution. This mechanism is implemented associating pheromone levels with each arc, which are then updated proportionally to the goodness of the solutions found. In such a way the pheromone levels encode locally the global information on the solution. Each artificial ant will then use this information weighted with an appropriate local heuristic function (e.g. for the TSP, the inverse of the distance) during the construction of a solution.

Solving the VRP is known to be a combinatorial NP-hard optimization problem. When an exact approach exists, it often requires large computational times, and is not viable in the time scale of hours, usually the time scale required by distribution planners. With the development of real-time data acquisition systems, and the consideration of various dynamic aspects, it appears more and more advisable to find high quality solutions to updated information in sensibly shorter times. In this study which is different with above four ACA mainly in two aspects as follows: (1) to increase the local searching ability, dynamic Sweep Algorithm is applied in the ACA originally. (2)A new Multiple Ant Colonies technique, which is completely different from other,

is developed. Our computational findings show that a considerable improvement is achieved through this new approach.

4 Hybrid Ant Colony Algorithm Based on Dynamic Sweep Algorithm for VRPTW

4.1 Dynamic Sweep Algorithm

The dynamic sweep is imitates one reconnaissance method which in the military the radar scan, it use the general sweeping principle, dynamic chooses the sweeping beginning according to the system request and complete to the goal sweeping. This article uses this kind of principle, transports to the logistics distribution needs the customer the demand and the position carries on the sweeping, then allocates and dispatches the region according to the logistics center's resources reasonable division.

Dynamic sweep rules: From a logistics center to the point of the initial starting point-ray, anti-clockwise (or clockwise) scan, the sweeping process, each faced a client under current vehicles determine whether the capacity of the customer assigned to the current vehicles, if not contrary to capacity constraints; if no, on the re-use another car to reach this customer, using the formula (2) to judge. So repeatedly, until all of the customers is assigned various vehicles, as shown in Figure 1.

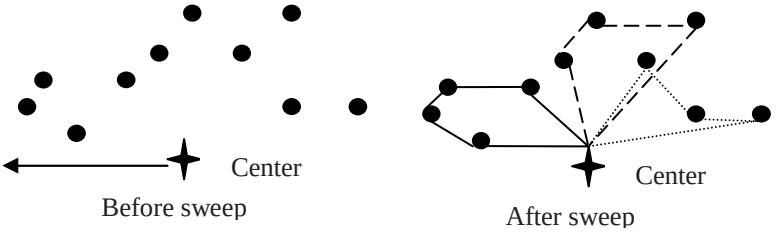


Fig. 1. Dynamic sweep example. The spot represent the customer, the cross star represents the logistics center; the different linear connection customer represents the different grouping.

4.2 Improved Ant Colonies Algorithm Design

In the logistics transportation, first used dynamic sweeping algorithm division of the customers, then used ant colony algorithm for each vehicle assigned to customers. At each iteration of the basic Ant Colony method, each ant builds a solution of the problem step by step. At each step the ant makes a move in order to complete the actual partial solution choosing between elements of a set A_k of expansion states, following a probability function. For each ant *k* probability $p_{ij}^{k}(t)$ of moving from present state i to another state j is calculated taking into account:

1. Attractiveness η_{ij} of the move according to the information of the problem.
2. Level τ_{ij} of pheromone of the move that indicates how good the move was in the past.

3. A allowed$_k$ list of forbidden moves; in the ant algorithm original version formula for $p_{ij}^k(t)$ is:

$$p_{ij}^k(t)=\begin{cases}\dfrac{[\tau_{ij}(t)]^{\alpha}[\eta_{ij}(t)]^{\beta}}{\sum_{s\in allowed_k}[\tau_{ij}(t)]^{\alpha}[\eta_{ij}(t)]^{\beta}} & j\in allowed_k\\ 0 & otherwise\end{cases}, \tag{10}$$

where α, β are parameters that are used to establish the relative influence of η_{ij} versus τ_{ij}. After iteration t is complete, that is when all the ants have completed their solutions, the pheromone levels are updated to:

$$\tau_{ij}(t+1)=\rho.\tau_{ij}(t)+\Delta\tau_{ij}\quad \Delta\tau_{ij}=\sum_{k=1}^{m}\Delta\tau_{ij}^k, \tag{11}$$

$$\Delta\tau_{ij}^{\ k}=\begin{cases}Q/L_k & if\ \ ant\ \ k\ \ pass\ \ by\ \ path\ \ ij\\ 0 & otherwise\end{cases}. \tag{12}$$

where Q is a constant that ants week cycle of the total amount of pheromone. L_k representing the Kth ant which walks path length in this cycle, it reflects the shortest path within the overall situation, the search system to improve convergence speed. Q、C、α、β、ρ experimental methods can be used determine their optimal combination.

4.2.1 Pheromone Updating

In the real world, the higher concentration of pheromone, volatile sooner, and the lower concentration of pheromone, the more diffuse volatile, so effective in preventing some pheromone concentration on the path to unlimited growth, and the pheromone on some path reduce to zero, reducing the possibility of a local optimum. In such circumstances volatile factor from a constant to a variable function by τ_{ij}, therefore, use the following update rules:

$$\tau_{ij}(t+N_{\max})=(1-\rho*\tau_{ij}(t))\cdot\tau_{ij}(t)+(\rho*\tau_{ij}(t))*\Delta\tau_{ij}(t,t+N_{\max}), \tag{13}$$

$$\Delta\tau_i(t,t+N_{\max})=\sum_{k=1}^{m}\Delta\tau_i^k(t,t+N_{\max})\quad \Delta\tau_i^k(t,t+N_{\max})=\frac{\mathrm{Q}}{\mathrm{L}_{kl}}. \tag{14}$$

where $N_{\max}$ is a once cycle required the maximum time. Assume a unit time before ant's progress; L_{kl} representing total distance traveled of vehicle routing *k*.

4.3 Mixed Ant Colony Algorithm Process

1. According to a logistics centre, customers and geographic information on the dynamic sweep path line. each customer with a polar coordinate, logistics centre is the origin o, either a node i is 0 degrees position, each customer j coordinate is(θ_j, δ_j), θ_j is angle, δ_j is the ray length, for all the customers follow by θ_j ascending.

2. Select unused vehicles *k*, beginning of logistics centre from the minimum angle designated for the vehicle *k* arrived, according θ_j to visit, ascending until the capacity is not satisfied with, accord (2) restrictions.
3. If all the customers were designated of the vehicle, stop sweeping; otherwise, repeat step 2 until all customers have arranged vehicles.
4. Test have unreasonable transferred customers because of the customers, makes the vehicle which redeployed to meet customer could not satisfied any of the time constraint, if found a number of customers unable to meet the arrival time, the group afresh start sweep.
5. Group After used dynamic sweeping; each group used ant algorithm for optimal. Each assigned vehicles, arrival visit order as follows. *K* vehicles assigned by the customer set for V_k and $|V_k| = m$, Defined only have *m* ants, definition $NC_{\max}$ is the largest cycles.

Step1 initialization: Set $t = 0$, $NC = 0$, on each edge $\tau_{ij}(0) = C$, and $\Delta\tau_{ij} = 0$, randomly placed ,*m* ant assigned to the *n* customer;

Step2 Order $s = 1$ (s is the tabu list subscript)

For $k = 1$ to *m* do, hold the Kth ant arrive customers number first to place $tabu_k(s)$;

Step3 repeat this step until the tabu list was filled (repeat this step $n-1$)

Set $s = s+1$ For $k = 1$ to *m* do

Under $p_{ij}^k(t)$ (of 10) to select the next step should be arrived customers *j*. at *t* moment ant *k* arrived at customer $i = tabu_k(s-1)$, moved to the Kth ants to customer *j*, and put *j* inserted to $tabu_k$ (s);

Step4 For $k = 1$ to *m* do

Put Kth ant from customers $tabu_k$ (*n*) to customers $tabu_k$ (1), calculating Kth ant for total path length, after update to find the shortest path.

For $k = 1$ to *m* do According to (14), updating edge of the pheromone $\Delta\tau_{ij}$;

Step5 According to (14), calculating each side $\tau_{ij}(t+n)$

Set $t = t+n$ Set $NC = NC+1$ Set $\tau_{ij} = 0$

Step6 if ($NC < NC_{MAX}$) and (not all ants choice the same path), then initialize all the tabu list, go to step2 first step; else, go to step7;

Step7 print out the shortest path, end the process, recorded the minimum value of the objective function and satisfied the time-bound solution.

5 Simulation Results

In order to justify the performance of the method developed in this study, DSACA was tested on Solomon standard data R102 given in [10]. These problems have larger calculation which simulation of the logistics transferring conditions. Data as shown in table 1, the specific data in [10]. R102 data have a total of 100 customers waiting for service, each vehicles volume is 200 units, the customers demand is less than the

Table 1. Logistics transferring

customers	coordinate	volume	time demand	service time
center	(35,35)	0	(0,230)	0
customer1	(41,49)	10	(0,204)	10
customer2	(35,17)	7	(0,202)	10
customer3	(55,45)	13	(0,197)	10
customer4	(55,20)	19	(149,159)	10
……	……	……	……	……
customer100	(18,18)	17	(185,195)	10

vehicles volume. Problem can be described as: if the customers were in100 different geographical location, and time to send have different requirements of the customer. All the vehicles service time for a customer is 10 units.

Use DSACA to solve the above problems, in matlab7.0, at the P-4 2.4G computing, initialization parameter settings $\alpha = 1$; $\beta = 1$; $\rho = 0.2$; $NC = 1000$. The final transportation optimization results: use 17 vehicles; the total distance 1503.8, the waiting time is 747.9. The vehicles arrangements are as follows: Vehicle 1 Path: 0, 63, 11, 19, 49, 7, 52, 0; Vehicle 2 Path: 0, 27, 69, 76, 68, 24, 80, 12, 0; Vehicle 3 Path: 0, 1, 33, 30, 51, 9, 66, 50, 0;Vehicle 4 Path: 0,92,37,98,91,16,86,85,93,59,0; Vehicle 5 Path: 0, 28, 29, 78, 34, 35, 3, 77; Vehicle 6 Path: 0, 21, 73, 22, 72, 54, 0; Vehicle 7 Path: 0, 48, 47, 82, 18, 0; Vehicle 8 Path: 0, 14, 44, 38, 43, 100, 95, 0; Vehicle 9 Path: 0, 65, 71, 81, 20, 32, 70, 0; Vehicle 10 Path: 0, 87, 57, 2, 0; Vehicle 11 Path: 0, 83, 45, 61, 84, 60, 89, 0; Vehicle 12 Path: 0, 40, 53, 0; Vehicle 13 Path: 0, 42,15,41,75,56,74,58, 0; Vehicle 14 Path: 0, 62, 88, 8, 46, 17, 5; Vehicle 15 Path: 0, 36, 64, 90, 10, 31, 0;Vehicle 16 Path: 0, 96,99,6, 94, 97, 13, 0; Vehicle 17 Path: 0, 39,23,67,55,4,25,26,0.

This paper proposed by the algorithm based on dynamic sweeping and ant colony algorithm (DSACA) which is able to find satisfactory solution.

6 Conclusions

In this paper, a hybrid ant colony algorithm for the vehicle routing with time windows (DSACA-VRPTW) is proposed. Although DSACA basically uses the concepts developed in previous Ant Colony Algorithms, it has two original contributions: the dynamic sweep algorithm and the multiple ant colonies technique. The first use dynamic sweeping algorithm for grouping, and then improve the use of ant colony algorithm to optimize each path. We effectively improve the optimization search speed, and ensure optimal performance of the overall situation. The simulation results show that the method of logistics transportation is the effectiveness and practicality.

Acknowledgments. This work is partially support by The 11th Five Years Key Programs for Science and Technology Development of China (No.2008BADA8B03, No. 2006BAD08B01), Henan Program for New Century Excellent Talents in University (No. 2006HANCET-15).

References

1. Clark, G., Wright, J.W.: Scheduling of Vehicles from a Central Depot to a Number of Delivery Point. Operations Research. 12, 568--581 (1964)
2. Dorigo, M., Maniezzo, V., Colorni, A.: The Ant System: Optimization by a Colony of Cooperating Agen. IEEE Transactions on Systems, Man and Cybernetics, Part B. 26 (1), 29--41 (1996)
3. Dorigo, M., DiCaro, G., Gambardella, L.M.: Ant Algorithms for Discrete Optimization. Artificial Life. 5(2), 137--172 (1999)
4. Bullnheimer, B., Hartl, R.F., Strauss, C.: Applying the Ant System to the Vehicle Routing Problem. In: Osman, I.H., Voss, S., Martello, S., Roucairol C.(Eds.), Meta-Heuristics: Advances and Trends in Local Search Paradigms for Optimization, Kluwer Academics, pp. 109--120 (1998)
5. Colorni, A., Dorigo, M., Maniezzo, V., Trubian, M.: Ant System for Job-shop Scheduling. Belgian Journal of Operations Research, Statistics and Computer Science (JORBEL). 34, 39--53 (1994)
6. Costa, D., Hertz, A.: Ants Can Colour Graphs. Journal of the Operational Research Society. 48, 295--305 (1997)
7. Dorigo, M., Gambardella L.M.: Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem. IEEE Transactions on Evolutionary Computation.1, 53--66 (1997)
8. Gambardella, L.M., Dorigo, M.: HAS-SOP: An Ant Colony System Hybridized with a New Local Search for the Sequential Ordering Problem, INFORMS. Journal on Computing. 12 (3), 237--255 (2000)
9. Gambardella, L.M., Taillard, E.D., Dorigo, M.: Ant Colonies for the Quadratic Assignment Problem. Journal of the Operational Research Society. 50, 167--176 (1999)
10. Solomon M.M.: Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints. Operations Research. 35,254--265 (1987)

A Biological Intelligent Access Control System Based on DSP and NIR Technology

Xingming Zhang, Yingshan Li, Zihao Pan, Wenjin Gu, and Jianfu Chen

School of Computer Science and Engineering, South China University of Technology
381#, Wushan road,Guang zhou,Guangdong, China, 510640
cszxm@scut.edu.cn uarta@126.com txjchen@gmail.com

Abstract. Face recognition, fingerprint recognition and other biometric technology are used widely now, and requirement of safety and reliability of Access Control in the market is getting higher and higher. The paper designs and implements Intelligent Biometric Access Control System based on DSP and NIR technology. It mostly describes the overall design, embedded system design, near infrared camera design, face recognition and verification algorithm design in server. By using NIR technology the system fundamentally change the impact of illumination on face recognition and resolve the limitations of face recognition in applications.

Keywords: DSP, Near Infrared, Biometric, Face Recognition, Fingerprint Recognition, Phone Number Recognition.

1 Introduction

With development and mature of biometric technology, face recognition, fingerprint recognition and other biometric recognition has been applied to access control system and provide for more secure protection and convenient management[1]. Face and fingerprint recognition access control system can improve the secure level and truly realize room management based on individual. Based on biological characteristics of face and fingerprint, with cell phone message identification, access control system ensures the reliability and ease of use.

Face recognition has been widely used in access control, attendance, the criminal investigation of police detection, video surveillance, and many other areas [2] [3]. It is playing more and more important role. But illumination invariant is the main bottleneck in face recognition. On the uneven illumination conditions, the recognition rate dropped rapidly, affecting the technology to further promotion [4] [5]. The improvement on software algorithm is a “temporary solution” for face recognition in illumination invariant. To resolve the root problem of face recognition lies in improvements of front acquisition-hardware.

This paper achieves embedded system based on DSP and biometric intelligent access control system which realizes face recognition, fingerprint recognition and cell phone texts message recognition.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 55 – 62, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2 System Architecture

We research and develop an intelligent embedded system as a client, completing facial image acquisition, cell phone number collection, fingerprint acquisition and identification and other front functionality. The server will be used to complete face recognition, face identification, SMS identification and record log of going in and out. Fingerprint recognition uses ZEM100 of ZhongKong Technology Company.

The system is based on LAN. The server saves information, connects device through serial port to send message. The workstation is a platform for system administrator for monitoring the state of access control. The intelligent embedded system connects the door locks and alarm, controls the door and collects the state of door. It collects face image when somebody wants to enter, fingerprint and password received by cell phone text message, and decides whether to open the door. Fig. 1 shows the structure of the system.

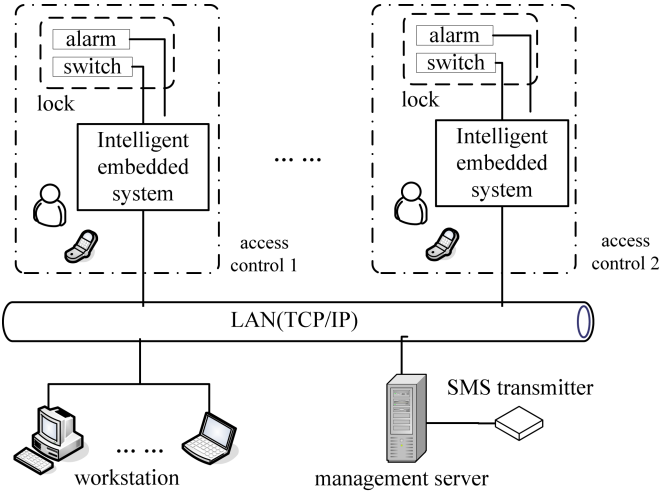


Fig. 1. Structure of the system

3 The Design of Embedded System

3.1 The Framework of Embedded Software

DSP/BIOS is an embedded, real-time, multi-task operating system which supports the management and deployment of multi-threading, includes three communication methods, Mailboxes, Semaphores and Queues, supports periodic function can be achieved timing function. The embedded of software access control system developed based on DSP/BIOS system, the framework is shown in the Fig. 2.

The embedded of software access control system includes three parts, I/O control part, video process part and communication part. The detailed functions are described as follows. (1)The I/O control part includes the operation of digital input and digital

output, the reading and writing of the serial interface and the process of interrupt service. The design of keyboard function is based on interrupt service and digital I/O. (2)The video process part includes the video capture, video compression and video transmission function. The detailed introduction will be described in the following chapter. (3)The communication part implements the communication between the access control client and the server. The module is developed based on NDK (Network Development Kit) of Texas Instruments [6]. The wireless communication module is designed for the better scalability.

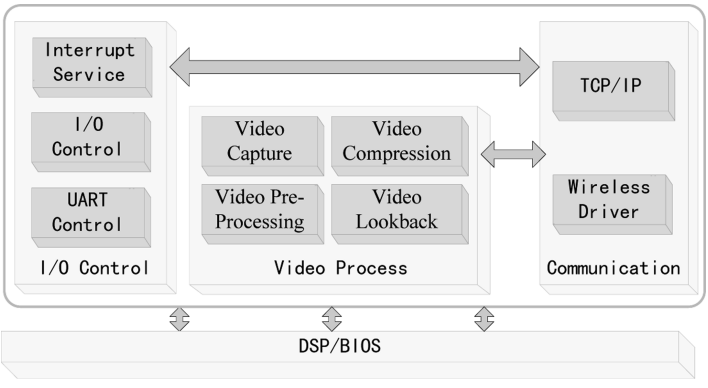


Fig. 2. The framework of embedded software

3.2 Data Flow Design

The system is designed based on RF5 (Reference Framework). RF5 contains the Memory Management module, Device Driver module and multithreading module, so the developer could focus on applications [7]. The data flow of the system includes three parts, the first is the video signal data flow encoded with JPEG arithmetic, the second is the image signal data flow saved with BMP format and the thirst is the control signal data flow.

Video signal data flow. GIO/FVID (the top layer class driver) transfer the video data to the buffer of the task receives the video. The task of video input gets the data package from the buffer and transfers the data format from YUV4:2:2 to YUV4:2:0. After finishing the transfer the video input task set the message package in which the address of the data is saved in the message queue. Then the video task waits the reply from video process task, while it gets the reply message it tells the class driver module and makes the buffer empty. At the same time the new video data will be got. The task of video encoding get the message of notifying process and get the video data from the input buffer named inBuf.

Image signal data flow. The face recognition of the access control system uses the BMP image sent from the client. The data flow of image signal is same as the data flow of video signal and the difference is it does not be compressed. In the main control task there is a switch which decides whether to compress or not. When the

server begins a recognition flow it sends the request to the client through the TCP/IP Socket. The network task gets the request and sends the command to the main control task through MailBox. The main control task notify the video process task, then the video process task send the data to the network task and the data will be sent to the server by the network task.

Control signal data flow. The main control task gets outer command through the interrupt service. If the command is keyboard interrupt the task will get keyboard value from I/O interface, if the command is serial interrupt the task will read the finger data from UART interface and if the command is the recognition finishing command the task will recall the function of opening the door. The control signal is transferred through MailBox. The main control task communicates with the receiving thread and the sending thread of the network task.

4 The Design of NIR Camera

4.1 The Theory of NIR Camera

The NIR camera is similar with the common camera and the key difference is the NIR camera uses the NIR light. In face these lights is LED with the high wavelength between 780nm to 1100nm. These LED are set around the camera and these lights and the camera must have the same direction. There is an optical filter which stops the visible light between CCD and the lens.

As is shown in the Fig. 3, the LED emits the NIR light to the face of user then the face reflects the light. The camera captures the light through the lens and the light is filtered by optical filter. At last the light is sent to the CCD and is transferred to the visible picture through spectral conversion and brightness enhance.

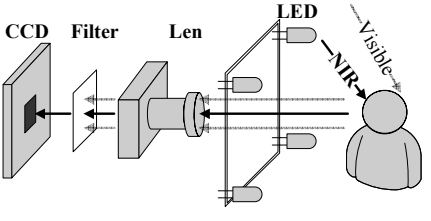


Fig. 3. Structure of NIR camera

The design of NIR Camera has several key points. The first is setting the optical filter, the filter must be set between the CCD and lens and it should be fixed on the CCD so that there is enough space for justifying focus of the camera. The second is the setting of NIR light, the lights should be placed around the lens symmetrically and keep the same direction with the CCD so that the face could get even light. The last is justifying the lens. After adding the filter the primary focus will change, so we must justify the focus of NIR camera through rotate the lens [8]. The selection of CCD is very important and we choose the bigger CCD and the monochrome CCD.

4.2 The Design of Filter

The filter is the core part of NIR camera. In the spectrum picture we could find that there are two areas in which the spectral radiometer capacity is very weak, one is the area with the wavelength from 925nm to 955nm and the other is from 845nm to 855nm. We choose 850nm as the main wavelength. We select the band pass filter of 850nm and it only allows the light between 845nm to 855 nm to pass. But the actual filter has some small pass area on both sides because of the limited technology. So we choose another long wavelength pass filter which could stop the light below 800nm. After grinding, bonding and focusing the near-infrared filter is formed.

4.3 The Application of Face Recognition

The use of NIR camera basically eliminates the influence of illumination changes. We obtain the photos with even illumination on the environment with uneven illumination. The photos in the first line of Fig.4 are captured when the user does not wear glasses and the photos in the last line are captured when the user wear glasses.

Fig. 4. The data flow of embedded software

Due to the reasonable distribution of NIR LED, the photos of the user wearing glasses are not influenced by the reflection of Near-infrared light on the glasses.

5 The Algorithm for Face Recognition and Verification

5.1 Face Recognition and Identification

In this chapter, we will propose “Auto Face Recognition Strategic”, which means how to automatically figure out whether the people’s face information has been registered or not. Then we will briefly introduce the face recognition strategic which bases on the Gabor Wavelet and the face identification strategic which bases on the Gabor Wavelet and the Support Vector Machine (SVM) [9].

Gabor wavelet 2-dimensional function is similar to the vision system of mammals. If we apply it to the face recognition, we will ideal result and higher robust for the light intensity transformation. We use the same method and parameters as [10] to extract Gabor feature.

5.2 Strategy of Automatic Face Recognition and Identification

Purpose and Processing. The Access Control System will use a time-sharing purview management mode in its daily running. In order to confirm whether the coming person has the purview, we will use a "Automatically Face Recognition Strategic", that means you can post a request without announcing any personal information , and our control system can automatically find out whether you have been registered or not. The work flow of this strategic is shown in Fig 5.

"Automatically Face Recognition Strategic" includes 2 parts. First, we will use the face recognition procedure to pick up some similar face information of the request petitioner's from the face database and package that information in a set. Then we will use the face identification procedure to validate the petitioner's face from the set one by one. If the result fits the standard, the petitioner will acquire the access purview.

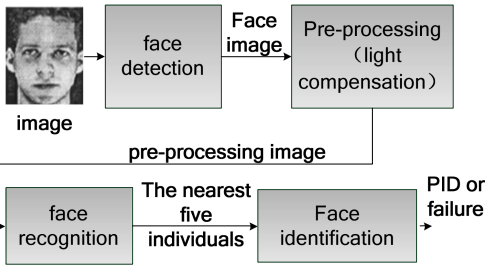


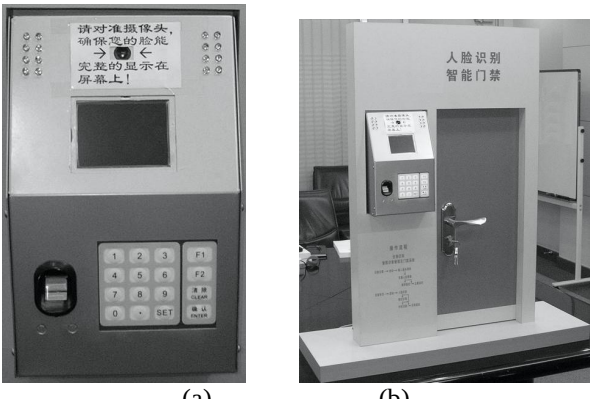
Fig. 5. The flow of automatically face recognition

Face Identification Strategy. In this part, we will use the following recognition strategic based on the Gabor feature. First, we will use the ATICR illumination treatment to transform the face image, Second we will use the Gabor transformation algorithm to extract the global and local (eyes, nose and mouth) face features and we will get a high dimensional features data. Third, we need to drop the dimensional space and put the result to the classification container to operate the preliminary face recognition. Finally, we will use the "Choquet Fuzzy Integral" algorithm to process the fusion recognition for the preview result and receive the final outcome [11].

Face Verification Strategy. Some professors propose a parallel wavelet approach for the face identification. First of all, they will use 40 Gabor kernel function to extract the face features, then use the PCA algorithm to drop the dimensional space of the Gabor features , and use the SVM to construct 40 classification containers, and finally use the additional principle to operate fusion recognition. But obviously this method requires too much data information and costs too much time. In the further research, we find that some wavelet functions don't have an ideal effect in the face identification, and have a high error access rate. After some test, we find that what we need to do is to pick up 16 SVM classification containers which have a higher recognition rate and lower error identification rate for the wavelet transformation and the result hit our demand.

6 Applications and Testing

The system had been built and set up on a simulative door. As is shown in Fig.6-(a), the NIR camera is set up on the top of the machine and the NIR LED is distributed on the both sides symmetrically.



(a) (b)

Fig. 6. Application systems. (a) The client machine of access control. (b) A simulative door.

Face is chief recognition method and we have a test using the images captured by NIR camera. It is a database with 50 persons with more than 50 images per person under one condition. And there are three different illuminating conditions, A, B, C. Three conditions are all in the room and condition A has the best illumination, condition C has the worst illumination. An average time of one person pass the face recognition is about 3.2 seconds. In Table 1, we can see that NIR images have a good performance in three conditions. Yet it is different between condition A and condition C, so we must improve the NIR camera and the algorithm for better application.

Table 1. Result of test

	A	B	C
FAR (%)	0	0	0.01
FRR (%)	0	0.25	2.25

7 Summary and Outlook

Based on TMS320DM642 and realizing face, fingerprint, and SMS identification, the intelligent embedded system takes full advantage of DM642 powerful video processing capability. The system realizes face recognition processing through network transmission of image, fingerprint recognition processing through serial communication, and SMS recognition. The innovation of the system using SMS recognition makes the system reliable and easy to use. NIR technology makes face recognition no

longer subject to the impact of illumination invariant, which greatly enhances the availability of the access control system. The proposed strategy of automatic face recognition makes the operating speed more useful.

The system is a more advanced and practical access control system and provides more security choices for users.

References

1. Yu, M., Zhen, H.: Face Recognition System Based on DSP. Microcomputer Information Press 19(6), 55 (2003)
2. Alex, P., Tanzeem, C.: Face Recognition for Smart Environments. IEEE Computer 2, 50–55 (2000)
3. Phillips, P.J., McCabe, R.M., Chellappa, R.: Biometric Image Processing and Recognition. In: Proceedings, European Signal Processing Conference (1998)
4. Dowdall, J., Pavlidis, I., Bebis, G.: Face Detection in the Near-IR Spectrum. Image and Vision Computing 21, 565–578 (2003)
5. Kong, S.G., Heo, J., Abidi, B., Paik, J., Abidi, M.: Recent Advances in Visual and Infrared Face Recognition - A Review. Computer Vision and Image Understanding 97(1), 103–135 (2005)
6. Li, S.Z., Zhang, L., Liao, S.C., Zhu, X.X., Chu, R.F.: A Near-Infrared Image Based Face Recognition System. In: Proc.Seventh IEEE Int'l Conf. Automatic Face and Gesture Recognition, April 2006, pp. 455–460 (2006)
7. Chen, X., Flynn, P.J., Bowyer, K.W.: Infra-Red and Visible-Light Face Recognition. Computer Vision and Image Understanding 99, 332–358 (2005)
8. Zhao, S.Y., Grigat, R.R.: An Automatic Face Recognition System in the Near Infrared Spectrum. In: Perner, P., Imiya, A. (eds.) MLDM 2005. LNCS (LNAI), vol. 3587, pp. 437–444. Springer, Heidelberg (2005)
9. Abhijit, Pandya, S., Macy, R.B.: Pattern Recognition and Implement in Neural Network. Publishing House of Electronic Industry
10. Liu, D.H., Lam, K.M., Shen, L.S.: Optimal sampling of Gabor features for face recognition. Pattern Recognition Letters 25, 267–276 (2004)
11. Gilbert, J.M., Yang, W.: A Real-Time Face Recognition System Using Custom VLSI Hardware. In: Proc. of Computer Architectures for Machine Perceptron Workshop, New Orleans, USA, pp. 58–66 (1993)

An Improved Support Vector Machine for the Classification of Imbalanced Biological Datasets

Haiying Wang and Huiru Zheng

School of Computing and Mathematics, University of Ulster, Shore Road, Newtownabbey, Co. Antrim, BT37 0QB,UK
{hy.wang,h.zheng}@ulster.ac.uk

Abstract. Based on advances in statistical learning theory, Support Vector Machine (SVM) has demonstrated unique features and state-of-the-art performance in many real-world classification problems. However, conventional SVM utilizes a sign function to classify test data into different classes, which has shown some limitations that hinder its performance. This paper explores the feasibility of incorporating information theory-based approaches into SVM decision making process and demonstrated its application in the classification of imbalanced biological datasets. The results obtained indicated that by incorporating information theory-based technique, a significant improvement was achieved ($p < 0.005$), especially in the process of classification of imbalanced datasets. The proposed methodology not only can improve the overall prediction performance but also can make the classification with the SVM less sensitive to the selection of input parameters.

Keywords: Support Vector Machine, Information theory, classification, Mouse Retina.

1 Introduction

Based on statistical learning theory presented in early 90's [1], support vector machine (SVM) has many unique features that make them attractive for classification and pattern recognition across a number of applications such as handwritten digit recognition [2], face detection in images [3] and information extraction [4]. In Bioinformatics, SVMs have been successfully applied to many areas, such as the analysis of microarray gene expression data [7], the recognition of translation initiation sites in DNA [5], and the identification of biomarkers in cancer research [6].

While demonstrating significant advantages in the process of classification in many complex domains, the SVM is not without its own limitations. For example, the conventional SVM always utilizes a sign function to classify test data into different classes. Zero becomes the only significant threshold value. Moreover, the whole prediction process doesn't take the decision value itself into account [11]. Such an approach may significant hinder its performance. For example, it has been shown that simply using sign function to make decision tends to produce a high rate of misclassification, especially for those points close to hyperplanes and for the case where the dataset is highly skewed [12].

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 63 – 70, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Different techniques have been proposed to improve decision making in the SVM. For example, Xie *et al.* [11] proposed a fuzzy SVM, which incorporates the decision value into prediction process. It not only gave the decision class for a given test sample but also indicated the membership to each class using the corresponding decision value. More recently, Li *et al.* [12] introduced a soft decision-making boundary to replace the conventional one with the sign function in the hope of improving the classification accuracy using the SVM.

This paper aims to propose the incorporation of information theory-based techniques to improve decision making in the SVM and demonstrates its application in the classification of imbalanced biological datasets. Based on the estimation of decision values associated with each training sample, a set of the best thresholds is identified and then a class label for each unseen case is predicted. The remainder of this paper is organized as follows. Section 2 overviews the basic concepts of SVM. Section 3 presents the methodologies used in this paper, followed by a brief description of the dataset under study. The results and the comparative analysis are given in Section 4. The paper concludes with discussion of the results, the limitations of this study and future research.

2 SVM: An Overview of Its Basic Concepts

The main idea behind the SVM is to construct a maximal separating hyperplane as the decision-making surface in an attempt to minimize empirical classification errors.

Let x_i be the input vector representing the i^{th} k-dimensional input sample and $y_i \in \{+1,-1\}$ be the class label indicating the class to which x_i belongs. In the case of linearly separable classification problems, the SVM algorithm simply needs to find a separating boundary which can be as far away from the data of two classes as possible. In the primal space, this can be formulated as the following constrained optimization problem:

$$\begin{aligned} &\textbf{Minimize: } \frac{1}{2}\|w\|^2 \\ &\textbf{Subject to: } y_i(w^T x_i + b) \geq 1, \quad \forall i \end{aligned} \tag{1}$$

where the vector w is normal to the heperplane and the value of $|b|/\|w\|$ decides the perpendicular distance from the hyperplane to the origin.

However, in reality, many problems are linearly non-separable. For example, biomedical data are characterized by the presence of large amounts of high-dimensional features whose dependence is usually an unknown nonlinear function [8]. It has been found that the solution derived from the linear SVM algorithm may not work well with non linear cases [9]. To deal with this problem, different methodologies have been used. For example, by using kernel trick, the original nonlinear problem can be mapped into a higher-dimensional (possibly infinite dimensional) feature space where the linear SVM can be used [10]. The soft margin-based technique introduces positive *slack variables* $\xi_i, i = 1,2,\cdots,n$ into the optimization problem [2], which then becomes:

$$\begin{aligned} &\textbf{Minimize: } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \\ &\textbf{Subject to: } y_i(w^T x_i + b) \geq 1-\xi_i, \ \ \xi_i \geq 0, \ \ \forall i \end{aligned} \tag{2}$$

The parameter C, representing the degree of tolerance, is used to control the trade-off between the error and margin.

Conventionally, performing classification with SVM involves the following steps.

- Training process. Using a given training dataset, $\{x_i, y_i\}$, $i = 1,2,\cdots,n,$, to decide the parameters for constructing a SVM-based classification model, for example, the type of kernel function used and the value of the cost parameter C.
- Identifying a set of support vectors, $s_i, i = \{1,2,\cdots,k\}$, which are located on the hyperphanes and constrain the width of the margin.
- Calculating the decision value, sv, for each test sample, x, using the following formula:

$$sv = \sum_{i=1}^{k} \alpha_i y_i K(x, x_i) + b \tag{3}$$

- Predicting the class of a test case based on the sign of its decision value, i.e.

$$y = \begin{cases} +1, \ if \ sv > 0 \\ -1, \ if \ sv < 0 \end{cases} \tag{4}$$

3 Methodology

The framework for the proposed information theory-based approach to improving decision making in the SVM is illustrated in Fig. 1. Like conventional SVMs, after the training process, a set of support vectors and the associated Lagrange multipliers were obtained. Thus, a model for calculating decision values for a given input sample is constructed. After that, we compute decision values for both training and testing data samples. Unlike the conventional SVM, which uses a sign function to determine the class label for a test sample, in this study, we incorporate information theory-based approaches into prediction process.

The idea is based on C4.5 learning algorithm [13]. Let $sv_i, i = 1,2,\cdots,n$, represent the decision values for the training data $x_i, i = 1,2,\cdots,n$. The associated label is indicated by $y_i \in \{+1,-1\}$. The entropy of the set SV is defined as

$$Entropy(X) = -p(pos)\log_2^{p(pos)} - p(neg)\log_2^{p(neg)} \tag{5}$$

where $p(pos)$ is the proportion of positive cases ($y_i = +1$) in the training data and $p(neg)$ is the proportion of negative cases ($y_i = -1$) . The scale of entropy reflects the class distribution. The more uniform the distribution, the smaller the calculated entropy and thus the greater the information obtained. Obviously, $Entropy(SV)$ is set to

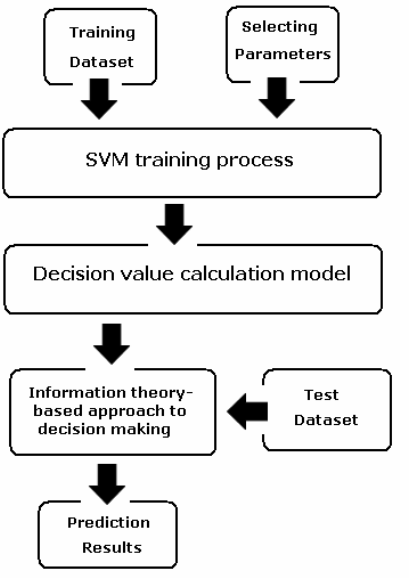


Fig. 1. The framework for the incorporation of the information theory-based approach to decision making in SVM

zero when all training members belong to the same class. In order to find a set of the best thresholds that can be used to make prediction for test data, the *information gain*, which is defined as the expected reduction in entropy caused by partitioning the data according to decision values, is calculated:

$$Gain(X, SV) = Entropy(X) - \sum_{v \in Value(SV)} \frac{|X_v|}{|X|} Entropy(X_v) \tag{6}$$

where *Value*(*SV*) is the set of all possible decision values calculated for training data, X, and X_v is the subset of X whose decision value is equal to v, i.e., $X_v = \{x \in X \mid SV(x) = v\}$. It has been shown that the information gain obtained from Equation (6) can be used to assess the effectiveness of an attribute in classifying training data and by maximizing the computed information gain, a set of the best thresholds, which is used to classify unseen samples into relevant classes, can be identified [13]. This process is similar to C4.5 decision tree-based classification analysis.

3 The Dataset under Study

The proposed methodology was tested on a biological dataset, which was constructed from a study published by Blackshaw *et al.* [14]. This database comprises a total of 14 murine gene expression libraries generated from different tissues and mouse developmental stages, including mouse 3T3 fibroblast cells, adult hypothalamus, developing retina at 2 day intervals from embryonic day (E) 12.5 to postnatal day (P) 6.5, P10.5

retinas from the paired-homeodomain gene crx knockout mouse (crx-/-) and from wild type (crx+/+) littermates, adult retina and microdissected outer nuclear layer (ONL). *Serial analysis of gene expression* (SAGE)-based technique, which allows simultaneous analysis of thousands of transcripts without *prior* knowledge of the gene sequences, was used to generate expression data. A total of 50000 - 60000 SAGE tags were sequenced from each tissue library. A detailed description of the generation and biological meaning of these libraries can be found in [14].

The problem posed in this dataset is to decide whether a tag represents a photoreceptor(PR)-enriched or a non-PR-enriched gene given a set of 14 gene expression libraries (i.e. 3t3, hypo, E12.5, E14.5, E16.5, E18.5, P0.5, P2.5, P4.5, P6.5, P10.5Crx-/-, P10.5Crx+/+, Adult, and ONL) associated with each tag. In order to control for sampling variability and to allow expression examination via *in situ* hybridization (ISH), we focused on those tags whose abundance levels represent at least 0.01% of the total mRNA expression in the ONL library. In this study, the training dataset includes 64 known PR-enriched genes and 10 non-PR-enriched genes validated prior to Blackshaw *et al*. 's study, while the test dataset consists of 197 PR-enriched genes and 53 non-PR-enriched genes, which were validated by Blackshaw *et al.*'s study using ISH [14]. As can be seen, there is a significant disproportion in the number of SAGE tags belonging to two functional classes. The difficulties posted in this dataset were further highlighted by the fact that patterns exhibited by PR-enriched genes are quite complex and diverse [17].

4 Results

The proposed methdology was implemented within the framework provided by the open source Weka package [15]. The implementation of SVM is based on sequential minimal optimization algorithm developed by Platt [16]. Two kernel functions: polynomial and radial basis functions (RBF) were used.

We firstly used conventional SVM to perform the classification tasks. Given that the dataset is highly imbalanced (training dataset: 64 belonging to majority class and 10 belonging to minority class), the prediction results were assessed in terms of four statistical indicators: *overall classification accuracy* (*AC*), *precision* (*Pr*), *sensitivity* (*Se*) and *specificity* (*Sp*), as shown in Table 1. The performance of classifiers based on conventional SVM was heavily dependent on the selection of kernel functions and the cost parameter *C*. For example, classification with polynomial functions produced very low sensitivity for negative class, implying that it failed to distinguish minority class from majority class in this case. Similar observation was made when using RBF as the kernel function with the value of *C* less than 50.

The same dataset was used to train the SVM but with a different decision making strategy. As illustrated in Table 2, by incorporating information theory into decision-making process, better prediction results were achieved. A closer examination of the results reveals that the integration of information theory–based approach to decision making not only improves the overall performance but also make the prediction results are less sensitive to the selection of input parameters. For example, the overall classification accuracy is between 82.4% and 84.00% when using RBF function with the value of *C* changing from 1 to 100.

Table 1. Prediction results using conventional SVMs. Two kernel functions: RBF and polynomial were used. *C* is the cost parameter, representing the tradeoff between the training errors and rigid margins. The parameter of *d* indicates the exponential value of the polynomial function.

Kernel function	AC (%)	Positive Class			Negative Class		
		Pr (%)	*Se* (%)	*Sp* (%)	*Pr* (%)	*Se* (%)	Sp (%)
RBF (C =1.0)	78.8	78.8	100	0	0	0	100
RBF (C =10)	79.2	79.1	100	1.9	100	1.9	100
RBF (C =30)	79.2	79.6	99.0	5.7	60.0	5.7	99.0
RBF (C =50)	80.8	81.2	98.5	15.1	72.7	15.1	98.5
RBF (C =80)	83.2	83.8	97.5	30.2	76.2	30.2	97.5
RBF (C =100)	83.2	83.8	97.5	30.2	76.2	30.2	97.5
Polynominal (d = 3, C = *50*)	78.8	78.8	100	0	0	0	100
Polynominal (d = 3,C = *100*)	78.8	78.8	100	0	0	0	100

Table 2. Prediction results with information theory-based approaches to decision making in SVM. Two kernel functions: RBF and polynomial were used. *C* is the cost parameter, representing the tradeoff between the training errors and rigid margins. The parameter of *d* indicates the exponential value of the polynomial function.

Kernel function	AC (%)	Positive Class			Negative Class		
		Pr (%)	*Se* (%)	*Sp* (%)	*Pr* (%)	*Se* (%)	Sp (%)
RBF (C =1.0)	82.8	84.1	96.4	32.1	70.8	32.1	96.4
RBF (C =10)	83.2	84.8	95.9	35.8	70.4	35.8	95.9
RBF (C =30)	82.8	86.3	92.9	45.3	63.2	45.3	92.9
RBF (C =50)	84.0	84.9	97.0	35.8	76.0	35.8	97.0
RBF (C =80)	84.0	84.9	97.0	35.8	76.0	35.8	97.0
RBF (C =100)	84.0	84.9	97.0	35.8	76.0	35.8	97.0
Polynominal (d = 3, C = *50*)	82.4	85.3	93.9	39.6	63.6	39.6	93.9
Polynominal (d = 3,C = *100*)	83.2	87.4	91.9	50.9	62.8	50.9	91.9

The significance of the differences between the conventional SVM and the proposed information theory-based approach to decision making in terms of the overall accuracy and the sensitivity for the minority class is established by a two-tailed *t* test. The obtained results indicate that the differences between two approaches are highly significant ($p < 0.005$).

5 Discussion and Conclusions

In the current study we have investigated the feasibility of using information theory-based techniques to support decision making in the SVM and demonstrated its application in dealing with imbalanced classification problems in Bioinformatics. The results attained indicated that by incorporating the knowledge about the entropy and information gain into SVM decision making process, a significant improvement was achieved, especially in the process of classification of imbalanced datasets ($p < 0.005$). Moreover, the methodology not only can improve the overall prediction performance but also can make the classification with the SVM less sensitive to the selection of input parameters. Thus, it has the potential to improve prediction performance of SVM in the analysis of biological datasets.

In the current study, the proposed method was tested on a relatively small dataset. The behaviour of the methodology in the analysis of large biological datasets such as microarray gene expression data and protein-protein interaction networks deserves further investigation. Comparison with other relevant techniques such as the studies published by Li *et al.* [12] and Wang *et al.* [17] would also be part of future work.

Like other classification techniques, the models presented in this paper still need users to determine several parameters in advance such as the cost parameter *C*. The combination with other machine learning and statistical learning techniques, which may provide a better insight into the problem of automatic determination of the learning parameters, would be part of our future work.

This study contributes to the development of automated computational methods to improve biomedical knowledge discovery and data mining tasks.

References

1. Vapnik, V.: The Nature of Statistical Learning Theory, 2nd edn. Springer, Heidelberg (1999)
2. Cortes, C., Vapnik, V.: Support Vector Networks. Machine Learning 20, 273–297 (1995)
3. Osuna, E., Freund, R., Girosi, F.: Training Support Vector Machines: An Application to Face Detection. In: Conf. Computer Vision and Pattern Recognition, pp. 130–136 (1997)
4. Li, Y., Bontcheva, K., Cunningham, H.: Using uneven images SVM and perceptron for information extraction. In: the 9th Conference on Computational Natural Language Learning, pp. 72–79 (June 2005)
5. Zien, A., Ratsch, G., Mika, S., Scholkopf, B., Lengauer, T., Muller, K.-R.: Engineering Support Vector Machine Kernels that Recognize Translation Initiation Sites. BioInformatics 16(9), 799–807 (2000)
6. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene Selection for Cancer Classification using Support Vector Machines. Machine Learning 46(1/3), 389–422 (2002)

7. Brown, M., Grundy, W., Lin, D., Cristianini, N., Sugnet, C., Furey, T., Ares Jr., M., Haussler, D.: Knowledge-based Analysis of Microarray Gene Expression Data by Using Support Vector Machines. Proc. Natl. Acad. Sci. 97, 262–267 (2000)
8. Altman, R.B.: Challenge for Intelligent Systems in Biology. IEEE Intelligent Systems 40(2), 394–409 (2001)
9. Burges, C.: A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2, 121–167 (1998)
10. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A Training Algorithm for Optimal Margin Classifiers. In: Haussler, D. (ed.) 5th Annual ACM Workshop on Computational Learning Theory, Pittsburgh, PA, pp. 144–152. ACM Press, New York (1992)
11. Xie, Z., Hu, Q., Yu, D.: Fuzzy Output Support Vector Machines for Classification. In: The Proc. of international conference on advances in natural computation 2005, pp. 1190–1197 (2005)
12. Li, B., Hu, J., Hirasawa, K.: An Improved Support Vector Machine with Soft Decision-Making Boundary. In: the 26th IASTED International Conference on Artificial Intelligence and Application (2008)
13. Quinlan, J.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo (1993)
14. Blackshaw, S., Fraioli, R.E., Furukawa, T., Cepko, C.L.: Comprehensive Analysis of Photoreceptor Gene Expression and the Identification of Candidate Retinal Disease Genes. Cell 107, 579–589 (2001)
15. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques, 2nd edn. Morgan Kaufmann, San Francisco (2005)
16. Platt, J.: Fast Training of Support Vector Machines using Sequential Minimal Optimization. In: Schoelkopf, B., Burges, C., Smola, A. (eds.) Advances in Kernel Methods - Support Vector Learning. MIT Press, Cambridge (1998)
17. Wang, H., Zheng, H., Simpson, D., Zauaje, F.: Machine Learning Approaches to Supporting the Identification of Photoreceptor-enriched Genes Based on Expression Data. BMC Bioinformatics 7, 116 (2006)

Prediction of Protein Homo-oligomer Types with a Novel Approach of Glide Zoom Window Feature Extraction

Qi-Peng Li[1,2], Shao-Wu Zhang[1], and Quan Pan[1]

[1] School of Automation, Northwestern Polytechnical University, Xi'an, China
[2] School of Mechatronics, Northwestern Polytechnical University, Xi'an, China
Authors' address: Qi-Peng Li, School of Automation/School of Mechatronics, Northwestern Polytechnical University, 127 YouYi West Rd., Xi'an 710072, Shaanxi, China
liqipeng@nwpu.edu.cn

Abstract. A novel approach of glide zoom window feature extraction based on protein sequence was proposed for predicting protein homo-oligomers. Based on the concept of glide zoom window, three feature vectors of amino acids distance sum, amino acids mean distance and amino acids distribution, were extracted. A series of feature sets were constructed by combing these feature vectors with amino acids composition to form pseudo amino acid compositions (PseAAC). The support vector machine (SVM) was used as base classifier. The 73.19% total accuracy is arrived in jackknife test, which is 7.87% higher than that of conventional amino acid composition method. The results show that the novel feature extraction method of glide zoom window is effective and feasible, and the feature vectors extracted with this method may contain more protein structure information.

Keywords: Glide Zoom Window, Feature Extraction, Pseudo Amino Acid Compositions, Homo-oligomer, Support Vector Machine.

1 Introduction

Proteins play an essential role in nearly all cell functions such as composing cellular structure and promoting chemical reactions. Given a polypeptide chain, will it form a dimer, trimer, or any other oligomer? This is an important task, because the functions of proteins are closely related to their quaternary attributes [1, 2]. The subunit construction of many enzymes provides the structural basis for the regulation of their activities, and indispensable function for many important biological processes. Thus, in the protein universe, there are many different classes of subunit construction, such as monomer, dimer, trimer, tetramer, and so forth. Some special functions are realized only when protein molecules are formed in oligomers; e.g., GFAT, a molecular therapeutic target for type-2 diabetes, performs its special function when it is a dimer [3], some ion channels are formed by a tetramer [4], and some functionally very important membrane proteins are of pentamer [5, 6,7].

It is generally accepted that the amino acid sequence of most, not all, proteins contains all the information needed to fold the protein into its correct three-dimension

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 71 – 78, 2008.
© Springer-Verlag Berlin Heidelberg 2008

structure structure [8,9]. At the next level of protein organization, tertiary structures associate into quaternary structures. Quaternary structure refers to the number of polypeptide chains (subunits) involved in forming a protein and the spatial arrangement of its subunits. The concept of quaternary structure is derived from the fact that many proteins are composed of two or more subunits that associate through non-covalent interactions and, in some cases, disulfide bonds. The association of subunits depends upon the existence of complementary 'patches' on their surfaces. The patches are buried in the interfaces formed by the subunits, thus, play a role in both tertiary and quaternary structure. Jones and Thornton [10,11] used a series of parameters to characterize and predict protein–protein interfaces on the basis of patch analysis. This suggests that primary sequences contain quaternary structure information [12].

Garian [12] predicted homodimer and non-homodimer using decision-tree models and a feature extraction method (simple binning function), and found that protein sequences contain quaternary structure information. Chou and Cai [13] also researched this question with a pseudo-amino acid composition feature extraction method. Zhang [14] classified homodimers and non-homodomers using the feature extraction method of amino acid index auto-correlation functions, and Zhang [15] also predicted protein homo-oligomer types by pseudo amino acid composition. In this paper, we propose the concept of glide zoom window based on the protein sequence, and try to develop the novel feature extraction method incorporating sequence order effect. Three feature vectors, which are amino acids distance sum, amino acids mean distance and amino acids distribution , were extracted from the glide zoom window of protein sequence . This new feature extraction method is combined felicitously with a support vector machine [16, 17] to predict homodimers, homotrimers, homotetramers and homohexamers.

2 Materials and Methods

2.1 Database

The dataset1283 consists of 1283 homo-oligomeric protein sequences, 759 of which are homodimers (2EM), 105 homotrimers (3EM), 327 homotetramers (4EM) and 92 homohexamers (6EM). This dataset was obtained from SWISS-PROT database [18] and limited to the prokaryotic, cytosolic subset of homo-oligomers in order to eliminate membrane proteins and other specialized proteins.

2.2 The Methods of Feature Extraction

Suppose the dataset consists of N homo-oligomeric protein sequences. L^k represents the sequence length of the kth protein sequence. α_i represents the ith amino acid of the nature amino acid set AA, $AA = \{A, R, N, D, C, O, E, G, H, I, L, K, M, F, P, S, T, W, Y, V\}$. Here, f_i^k and l_i^k represent the first position and last position of α_i in the k-th protein sequence p^k , respectively.

Glide zoom window of every nature amino acid can be described as a segment sequence of one protein sequence, that is, the glide zoom window of one nature amino acid begins from the position where this amino acid appears firstly and ends at the position where this amino acid appears lastly among the whole protein sequence. There are twenty glide zoom windows for every protein sequence. For example, for the protein sequence 'MITRMSELFLRTLRDDP', the glide zoom window of amino acid M is 'MITRM', the glide zoom window of amino acid T is 'TRMSELFLRT', the glide zoom window of amino acid D is 'DD', and so on. If one nature amino acid does not appear in the protein sequence, the glide zoom window of this nature amino acid is empty. The position and the width of every glide zoom window are variable. Apparently, the glide zoom window contains some sequence order information.

We can use z_i^k to represent the glide zoom window of α_i in p^k, which is the segment sequence between f_i^k and l_i^k. L_i^k is defined as length of z_i^k. We first defined a position indicator $s_{i,j}^k$ as:

$$s_{i,j}^k = \begin{cases} 1 \text{ if } \alpha_i \text{ locates in } jth \text{ position of } z_i^k \\ 0 \text{ if } \alpha_i \text{ does not locate in } jth \text{ position of } z_i^k \end{cases} \quad (1)$$

To integrate more sequence order information, according to the concept of glide zoom window, three feature vectors of glide zoom window are extracted to predict homo-oligomers. The three feature vectors of glide zoom window are defined as follows:

1) Amino Acids Distance Sum Feature Vector

The amino acids distance sum feature vector of p^k is expressed as the following 20-D feature vector:

$$AADS^k = \left[\xi_1^k, \cdots, \xi_i^k, \cdots, \xi_{20}^k \right], \quad k = 1, \cdots, N \quad (2)$$

Here,

$$\xi_i^k = \sum_{j=1}^{L_i^k} s_{i,j}^k \times j \quad (3)$$

Conveniently, the feature set based on this amino acids distance sum approach can be marked as S.

2) Amino Acids Mean Distance Feature Vector

The amino acids mean distance feature vector of p^k is expressed as the following 20-D feature vector:

$$AAMD^k = \left[\eta_1^k, \cdots, \eta_i^k, \cdots, \eta_{20}^k \right], \quad k = 1, \cdots, N \quad (4)$$

Here,

$$\eta_i^k = \begin{cases} \frac{1}{L_i^k} \sum_{j=1}^{L_i^k} s_{i,j}^k \times j \ , \ \text{if} \ L_i^k \neq 0 \\ 0 \qquad\qquad , \ \text{if} \ L_i^k = 0 \end{cases} \tag{5}$$

Conveniently, the feature set based on this amino acids mean distance approach can be marked as M.

3) Amino Acids Distribution Feature Vector

The amino Acids distribution feature vector of p^k is expressed as the following 20-D feature vector:

$$AAD^k = \left[\rho_1^k, \cdots, \rho_i^k, \cdots, \rho_{20}^k \right], \quad k = 1, \cdots, N \tag{6}$$

Here,

$$\rho_i^k = \begin{cases} \frac{1}{L_i^k} \sum_{j=1}^{L_i^k} \left(s_{i,j}^k \times j - \eta_i^k \right)^2, \ \text{if} \ L_i^k = 0 \\ 0 \qquad\qquad\qquad , \ \text{if} \ L_i^k = 0 \end{cases} \tag{7}$$

Conveniently, the feature set based on this amino acids distribution approach can be marked as D.

2.3 Support Vector Machine

The SVM is a binary classification algorithm [17, 19]. The algorithm has found wide application in many fields [19], including bioinformatics applications such as recognition of translation start sites, protein remote homology detection, microarray gene expression analysis, functional classification of promoter regions, and peptide identification from mass spectrometry data.

SVM has many strategies [20],such as: one-versus-rest [21] or one-versus-all [22], one-versus-one [21] or all-versus-all [22], DAGSVM [23] and so on. Three typical kernel functions used in SVM are: Polynomial function, Radial basis function (RBF) and Sigmoid function. one-versus-one and Radial basis function (RBF) are more effective [20]. The SVM algorithm implemented here (LIBSVM) can be download freely (http://www.csie.ntu.edu.tw/~cjlin/libsvm/).

2.4 Assessment of the Prediction System

The prediction quality can be examined using the jackknife test. The cross-validation by jackknifing is thought the most objective and rigorous way in comparison with sub-sampling test or independent dataset test [24, 25]. During the process of jackknife analysis, the datasets are actually open, and a protein will in turn move from each to the other. The total prediction accuracy (Q) and Matthew's Correlation Coefficient

(MCC) [26] for each class of homo-oligomers calculated for assessment of the prediction system are given by:

$$Q = \sum_{i=1}^{M} p_i \Big/ N \times 100\% \tag{8}$$

$$Q(i) = p_i / (p_i + u_i) \tag{9}$$

$$MCC(i) = \frac{p_i n_i - u_i o_i}{\sqrt{(p_i + u_i)(p_i + o_i)(n_i + u_i)(n_i + o_i)}} \tag{10}$$

Here, N is the total number of sequences, p_i is the number of correctly predicted sequences of *k* class protein homo-oligomers, u_i is the number of under-predicted sequences of *k* class protein homo-oligomers, n_i is the number of correctly predicted sequences not of *k* class protein homo-oligomers, o_i is the number of over-predicted sequences of *k* class protein homo-oligomers.

3 Results and Discussion

3.1 The Results of Different Pseudo Amino Acids Composition Feature Sets

The feature set based on the amino acid composition approach [27] can be marked as C. Using the feature set C and three glide zoom window feature sets S, M and D, the seven feature sets of pseudo amino acid composition (PseAAC) are constructed. The results of these seven PseAAC feature sets and feature set C with RBF SVM and one-versus-one strategy in jackknife test are shown in table 1.

From Table 1, we can see that the results of CM is the best in all the feature sets, and the total accuracy is 74.59%, which is 5.77% higher than that of feature set C. These results suggest that, in glide zoom window , the feature set of amino acids mean distance is more effective and robust than other feature sets. In addition, the

Table 1. Results of 8 Feature sets with RBF SVM and one-versus-one strategy in jackknife test

Feature sets	2EM		3EM		4EM		6EM		Q%
	Q(2)%	MCC(2)	Q(3) %	MCC(3)	Q(4) %	MCC(4)	Q(6) %	MCC(6)	
C	91.57	0.3582	42.86	0.5726	38.53	0.3568	18.48	0.3088	68.82
CM	91.17	0.4943	53.33	0.6511	55.35	0.5053	30.43	0.4373	74.59
CS	94.33	0.3589	36.19	0.5150	37.61	0.3753	3.26	0.1439	68.59
CD	95.39	0.3302	32.38	0.5276	33.03	0.3611	1.09	0.0992	67.58
CMS	91.17	0.4963	53.33	0.6572	55.05	0.4973	31.52	0.4480	74.59
CMD	91.04	0.4916	53.33	0.6511	55.05	0.4989	30.43	0.4373	74.43
CSD	95.92	0.3281	28.57	0.4922	32.11	0.3569	1.09	0.0992	67.34
CMSD	90.78	0.4884	53.33	0.6634	55.35	0.4995	31.52	0.4480	74.43

accuracies of CS, CD and CSD are lower than that of feature set C. The reasons are that there may be some redundancy and conflict information between these feature sets, or the unbalance of sample numbers among the four classes.

3.2 The Influence of the Unbalance of Sample Numbers among the Four Classes

We used the weighted factor approach to investigate the influence of the sample unbalance among the four classes. According to the number of four types of protein homo-oligomer, the weighted factor values of 2EM, 3EM, 4EM and 6EM are calculated as follow: 759/759, 759/105, 759/327, 759/92. The results of 8 feature sets using weighted factor approach are shown in table 2.

From table 2, we can see that the total accuracies of the seven feature sets CM, CS, CD, CMS, CMD, CSD, CMSD are higher than that of feature set C in the weighted factor conditions. The result of CMSD is the best, and the total accuracy are 73.19%, which are 7.87% higher than that of feature set C. These results suggest that weighted factor approach can weaken influence of the unbalance of sample numbers among the four classes.

Table 2. Results of 8 feature sets with RBF SVM and one-versus-one strategy in jackknife test using weighted factor approach

Feature sets	2EM		3EM		4EM		6EM		Q%
	Q(2)%	MCC(2)	Q(3) %	MCC(3)	Q(4) %	MCC(4)	Q(6) %	MCC(6)	
C	70.36	0.3577	49.52	0.4772	63.91	0.3859	46.74	0.3752	65.32
CM	78.00	0.4647	59.05	0.5532	67.58	0.5035	53.26	0.5188	72.0
CS	76.81	0.4363	55.24	0.5371	66.36	0.4665	45.65	0.4305	70.15
CD	76.02	0.4105	53.33	0.5213	64.83	0.4383	42.39	0.4092	68.90
CMS	78.39	0.4735	59.05	0.5604	67.89	0.5025	53.26	0.5270	72.33
CMD	78.79	0.4723	60.00	0.5677	66.97	0.5039	54.35	0.5356	72.49
CSD	75.76	0.4271	57.14	0.5300	64.83	0.4537	46.74	0.4127	69.37
CMSD	80.24	0.4837	60.00	0.5866	66.36	0.5077	54.35	0.5443	73.19

4 Conclusion

In this paper, we proposed a novel concept of glide zoom window. Based on the concept of glide zoom window, a series of feature sets were constructed by combining three feature vectors of glide zoom window with amino acids composition to form pseudo amino acid compositions (PseAAC). The results show that the seven feature sets based on the novel feature extraction method of glide zoom window are better than feature set C in the weighted factor conditions, and weighted factor approach can weaken influence of the unbalance of sample numbers among the four classes. Amino acids mean distance feature set of glide zoom window is most effective and robust in the three feature sets of glide zoom window. It is demonstrated that the feature sets of glide zoom window may contain more protein structure information, and this novel feature extraction method can be used to predict other protein attributes.

Acknowledgements. This paper was supported in part by the National Natural Science Foundation of China (No. 60775012 and 60634030) and the Technological Innovation Foundation of Northwestern Polytechnical University (No. KC02).

References

1. Chou, K.C.: Review: Low-frequency Collective Motion in Biomacromolecules and Its Biological Functions. Biophys. Chem. 30, 3–48 (1988)
2. Chou, K.C.: Review: Structural Bioinformatics and Its Impact to Biomedical Science. Curr. Med. Chem. 11, 2105–2134 (2004e)
3. Chou, K.C.: Molecular Therapeutic Rarget for Type-2 Diabetes. J. Proteome Res. 3, 1284–1288 (2004a)
4. Chou, K.C.: Insights from Modeling Three-dimensional Structures of the Human Potassium and Sodium Channels. J. Proteome Res. 3, 856–861 (2004)
5. Chou, K.C.: Insights from Modeling the 3D Structure of the Extracellular Domain of Alpha7 Nicotinic Acetylcholine Receptor. Biochem. Biophys. Res. Commun. 319, 433–438 (2004)
6. Chou, K.C.: Modelling Extracellular Domains of GABA-A Receptors: Subtypes 1, 2, 3, and 5. Biochem. Biophys. Res. Commun. 316, 636–642 (2004)
7. Oxenoid, K., Chou, J.J.: The Structure of Phospholamban Pentamer Reveals a Channel-like Architecture in Membranes. Proc. Natl. Acad. Sci. USA 102, 10870–10875 (2005)
8. Anfinsen, C.B., Haber, E., Sela, M., White, F.H.: The Kinetics of the Formation of Native Ribonuclease During Oxidation of the Reduced Polypeptide Chain. Proc. Natl. Acad. Sci. USA 47, 1309–1314 (1961)
9. Anfisen, C.B.: Principles That Govern the Folding of Protein Chains. Science 181, 223–230 (1973)
10. Jones, S., Thornton, J.M.: Analysis of Protein–protein Interaction Sites Using Surface Patches. J. Mol. Biol. 272, 121–132 (1997a)
11. Jones, S., Thornton, J.M.: Prediction of Protein–protein Interaction Sites Using Patch Analysis. J. Mol. Biol. 272, 133–143 (1997b)
12. Garian, R.: Prediction of Quaternary Structure from Primary Structure. Bioinformatics 17, 551–556 (2001)
13. Chou, K.C., Cai, Y.D.: Predicting Protein Quaternary Structure by Pseudo Amino Acid Composition. Proteins Struct. Func. Gene. 53, 282–289 (2003b)
14. Zhang, S.W., Quan, P., Zhang, H.C., Zhang, Y.L., Wang, H.Y.: Classification of Protein Quaternary Structure with Support Vector Machine. Bioinformatics 19, 2390–2396 (2003)
15. Zhang, S.W., Pan, Q., Zhang, H.C., Shao, Z.C., Shi, J.Y.: Prediction of Protein Homo-oligomer Types by Pseudo Amino Acid Composition: Approached with an Improved Feature Extraction and Naive Bayes Feature Fusion Amino Acids (2006)
16. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (1995)
17. Vapnik, V.: Statistical learning theory. Wiley, New York (1998)
18. Bairoch, A., Apweiler, R.: The SWISS-PROT Protein Data Bank and Its New Supplement TrEMBL. Nucleic Acids Res. 24, 21–25 (1996)
19. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines. Cambridge University Press, Cambridge (2000)
20. Hsu, C.W., Lin, C.J.: A Comparison of Methods for Multi-class Support Vector Machines. IEEE Transactions in Neural Networks 13(2), 415–425 (2002)
21. Kreßel, U.H.: Pairwise Classification and Support Vector Machines. In: Schölkopf, B., Burges, C.J., Smola, A.J. (eds.) Advances in Kernel Methods: Support Vector Learning: 1999, pp. 255–268. MIT Press, Cambridge (1999)

22. Ding, C.H., Dubchak, I.: Multi-class Protein Fold Recognition Using Support Vector Machines and Neural Networks. Bioinformatics 17(4), 349–358 (2001)
23. Platt, J., Cristianini, N., Shawe-Taylor, J.: Large Margin Dags for Multiclass Classification. In: Jordan, M.I., Lecun, Y., Solla, S. (eds.) Proceedings of Neural Information Processing Systems, pp. 547–553. MIT Press, Cambridge (2000)
24. Chou, K.C., Zhang, C.T.: Review: Prediction of Protein Structural Classes. Crit. Rev. Biochem. Mol. Biol. 30, 275–349 (1995)
25. Zhou, G.P.: Assa-Munt N Some Insights Into Protein Structural Class Prediction. Proteins Struct. Funct. Genet. 44, 57–59 (2001)
26. Fasman, G.D.: Handbook of Biochemistry and Molecular Biology, 3rd edn. CRC Press, Boca Raton (1976)
27. Bahar, I., Atilgan, A.R., Jernigan, R.L., Erman, B.: Understanding the Recognition of Protein Structural Classes by Amino Acid Composition. Proteins 29, 172–185 (1997)

Multi-layer Ensemble Classifiers on Protein Secondary Structure Prediction

Wei Li, Yuehui Chen, and Yaou Zhao

School of Information Science and Engineering
University of Jinan
Jinan 250022, P.R. China
liwei198348@163.com

Abstract. Protein secondary structure prediction is a bridge between amino acid sequence and tertiary prediction. Many methods have been used to improve the prediction accuracy and have a great development. Here a new method to predict the secondary structure is proposed by using classified database. The structure is similar with PSI-PRED method. But in the middle of the two neural network (NN) a neural ensemble is used to divided the initial database into two databases. Then the two databases are input into the next NN separately in order to predict the possible secondary structure. Experiments show that the proposed method achieves an average Q_3 score of 82% in CB396 database and 86% in CB513 database which produces the best prediction accuracy.

Keywords: Protein secondary structure prediction; Ensemble classification; PSSM.

1 Introduction

The prediction of protein structure has become one of the most important problems in molecular biology. The gap between the number of unknown structure of protein sequence and the known structure of the protein is growing faster and faster with the completion of the genome sequencing projects. Although the prediction of tertiary structure is one of the ultimate goals of protein science, the prediction of it is very difficult. So the prediction of the secondary prediction is an important middle step from sequence to structure [1].

Computational approaches to predict protein structures can be very useful in creating insights into protein folding and structure. All structure prediction methods basically rely on the idea that there is correlation between residue sequence and structure[3]. At the beginning of prediction, the most important method depends on homolog sequence of the known structure. But the method will fail when we can't find the homolog sequence or unable to identify homolog in the structural database. Then Chou and Fasman method and GOR method exploiting information from single residue and single sequence is proposed to predict protein secondary structure. Chou and Fashman method achieved only about 50%, and GOR is about 65%[2]. The GORV method has increased to

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 79–85, 2008.
© Springer-Verlag Berlin Heidelberg 2008

74.2% [1]. This method exploiting evolutionary information from multiple sequences which reveals the similarity of genetic code, evolutionary history, and common biological functions of the strings. But it is a challenge to detect proteins that share similar folds, but not clearly derived from a common ancestor [4]. PSSM (Position Specific Scoring Matrices) is a methods to get the derivation information from multiple sequence. Using PSSM information as the input of NN has achieved about 76% prediction accuracy. Neural networks are parallel, distributed information processing structure and the method is used to solve the problem by training the network. The most successful methods are PSI-PRED[5] and PHD[6,7].By now The higher prediction accuracy of method is PSI-PRED 76.0% with a standard deviation of 7.8%, and Byung-chul Lee's method is 79%[8]. The method present in this paper is a network method using neural ensemble to find out the uncertain residues in the protein set for predicting secondary protein structure.

There are 20 different amino acids in natural proteins, and there are eight types structure: helix (H), pi-helix (I), 3-10 helix (G), beta-strand (E), beta-bridge (B), h-bonded turn (T), bend (S), and coil (C). They are too many for existing methods to predict. instead, we usually use three states to represent the protein secondary structure: helix (H), extended (E), and coil (C). Here we used the definition of DSSP (database of secondary structure of protein). In the definition: H and G to helix (H), E and B to extend (E), and the others to coil (C).

The comparison of various methods for predicting secondary protein structure are three state per residue performances and Q_3 measure. CB396 and CB513 [4] databases are used and ten-folds cross validation is taken to test the performance of our method. When dataset of CB396 is used to evaluate the performance of the protein secondary structure prediction algorithms DSC, PHD, NNSSP and PREDATOR, the maximum Q_3accuracy is shown to be 78%.

In this paper , PSSM is presented in the section 1. The Algorithm is described in section 2.2. Section 3 gives the simulation results. Finally, we present some concluding remarks.

2 Methods

Our method is similar with the PSI-PRED method. Firstly, we input PSSM into the first NN in order to get the prediction output. Then the output are input into the second NN. The difference of them is that our method add a neural ensemble between the first and the second network. We introduce it in section 2.2. In section 2.1 PSSM is introduced.

2.1 PSSM

The PSSM is come from PSI-BLAST which used as input to the neural network.

It is based on the frequencies of each residue in a specific position of a multiple alignment. The times of one residue appearance in one place is the frequency, and then divided by the number of alignments.

In order to use it as the input of NN, a sliding window is applied along the sequence. The window width is 2×m+1, in the middle is the predict residue, and with m residues' value in both side of it. At the beginning and the end of the sequence there aren't m residues in the left or right of the middle residue, then use 0 to take it. In this paper, we take m equal to 5. So the window width is 11, and there are 20 types of residues. finally 20×11 values are used as the input of the NN.

The paper takes PSI-BLAST alignment with CB396 and CB513 in nr database to get PSSM. After removing the unknown structure residues and some wrong sequences, we get 61482 residues in CB396, 83628 residues in CB513. For ten-folds cross validation, the training data was divided into ten subdatabases with the same number of residues (the number of three different structure are equal in ten). With 9 sets as training and 1 set as test, we performed them for ten times.

2.2 Algorithm Description

The prediction of protein folding type is a typical multi-class data classification problem. Classification of multi-dimension data plays an important role in the decision to determine the main characteristics of a set. From the traditional methods it's easy to found that there are some "problematic instances" (PIs) influence the prediction. The PIs often have special character so that they are different from other individual and are difficult to be classified to one class. When training the NN with all the individuals, the PIs are not only difficult to be classified, but also may influence the excellent individual as yawp. Protein secondary prediction is a large project, as said before there are 60,000 to 80,000 individual to classify, and there would be too many PIs influence the result, for the prediction accuracy is lower than 80%. Selecting the PIs from the whole set and forming a new sets, then training the network, if a good result can be gotten?

We try to select the PIs from the initial database. The method can be divided into two steps: First, finding the PIs. The ensemble NNs are used here. The goal of the first stage is reveal the uncertainties in protein structure, and divided initial database into two different databases. Then, train the network separately to determine the three state probabilities for each residue.

Several different networks are used to predict the structure. Each of them finds out the incorrect prediction individuals. After that, several sets are produced, intersection set (IS) of them can be found out. Now all the PIs are gotten. Decompose the PIs is the good individuals. According to IS, primary set are divided into two sets called 'good' and 'bad'. The reason for keeping the 'bad' set is they also have residues' information of a sequence.

In this paper we use three separate layer as our predictor. The first NN is the main prediction neural,as the first layer; and the third one is for smoothing the forword network's results in order to account for the structural dependency between adjacent residues [8],which as the third layer.In the middle of them a

neural ensemble is used to divided the initial database into two databases, which as the second layer called the second NN.The algorithms is given as follows:

1. Input the initial datasets into the first NN. There are three outputs, and choose the biggest as the result of prediction. Supposing the order is H E C. For example, if output is 0.52, 0.12, 0.84, then choose the biggest as the result, then we think it was predicted to C.
2. INPUT:Similar to PSI-PRED, a sliding window is applied along the first NN's output. Here the window's width is 11. The second neural network's input is 11×3. The second NN's output is the same as the first NN.

   STRUCTURE:Five different networks by changing the learning rate, max iteration, neural number are presented. Input the 33 dimension individual into them. The incorrect prediction individual sets have been gotten.

   RESULT: Ensemble the five sets, if one individual appear in five sets, put it in the bad set. By now, we have divided initial set into two sets: 'good' and 'bad'.
3. Input 'good' and 'bad' into the third NN separately(the input is still the result of the first NN,the second NN only divided residue into two different databases). The output is the same as the first NN.Then the correct predicted individuals are adding together and divided the whole number of individuals to get the prediction accuracy. fig 1 give the description of the algorithm.

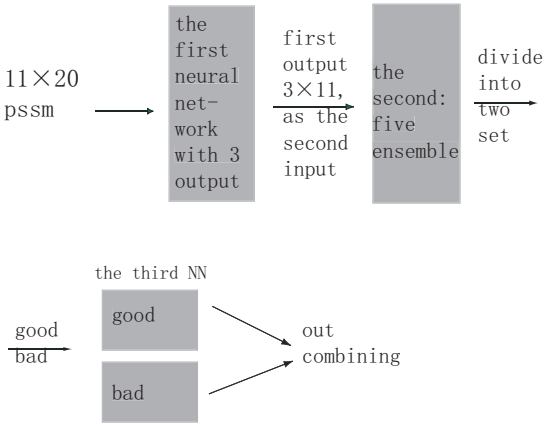


Fig. 1. PSSM is applied as the input of the NN. The second NN(neural ensemble)divided initial database into two,then input to the third NNs.

Table 1. PSI-PRED method on CB396 and CB513

	Q_3	Q_H	Q_E	Q_C
CB396	76.2	77.4	65.2	81.3
CB513	76.4	78.9	68.9	83.8

2.3 Standard Prediction Accuracy Measure

Traditionally, Q_3 measures the standard prediction accuracy [6]. Q_3 is a measure of the three-state overall percentage of prediction:

$$Q_3 = \frac{N_{correctlypredicted}}{N_{allresidues}} \times 100 \tag{1}$$

Per residue accuracy is defined below:

$$Q_3 = \frac{N^i_{correctlypredicted}}{N^i_{allresiduesofi}} \times 100, i\epsilon\{H, E, C\} \tag{2}$$

3 Result

It's often a difficult task to select important variables for a classification or regression problem, especially when the feature space is large. Conventional BP (Back Propagation Algorithm) neural network usually can't do this work. In our experiment leaning rate is 0.001, max generation is 50 to 200, the first layer's neural number is 30, the second is 10, which are better choosing for the experiment.

As mentioned before, the CB396 and CB513 database is used for cross validation. 9/10 of the samples is used for training the network and 1/10 for test. The process is repeated for 10 times. The result are shown in Table 1, when simulate the PSI-PRED.

The experiment is a control experiment that employs the conventional approach, which is using PSSM as input and three state as output. The PSI-PRED method's average Q_3 score is 76.0% with a standard deviation of 7.8%. Taken by residue, the average Q_3 score is 76.5%. Byung-chul Lee's method's [8] average Q_3 score is 78.5%,Q_H is 82.0%, Q_E is 66.4%, Q_C is 83.2%. Compared with them, it is easy to find that the control experiment can achieve the same effect with PSI-PRED.

In our research, the number of the network is 3. The second neural network is neural ensemble which find the individuals that are difficult to be classified. The third layer is the similar with the second. The difference between them is that for PSI-PRED the input is the initial dataset, while our method is "good" and "bad" datasets which have the same dimension. In the second neural network, we performed five networks to ensemble. The result of our method is shown in table 2.

From the result, we can find that our method can improve the prediction accuracy no less than 3%.

Experiment on the "good" and "bad" datasets, the average of ten-folds cross validation test result is shown in table 3 and table 4.

After decomposing the uncertainties individual, the prediction accuracy is improved evidently. it shows that the PIs greatly influence the prediction accuracy of secondary prediction.

Table 2. Ensemble Classification method on CB396 and CB513,compare with Byung-chul Lee's method

	Q_3	Q_H	Q_E	Q_C
CB396	82.2	77.4	75.6	89.8
CB513	86.5	90.5	70.9	91.6
Byung-chul Lee	78.5	82.0	66.4	83.2

Table 3. The result of good set

	Q_3	Q_H	Q_E	Q_C
CB396	88.6	83.9	84.9	93.1
CB513	91.8	95.4	83.6	93.2

Table 4. The result of bad dataset

	Q_3	Q_H	Q_E	Q_C
CB396	53.1	62.5	39.9	56.0
CB513	64.5	63.7	0.02	94.5

The prediction accuracy on 'bad' dataset is lower than 'good' dataset. We think that it's right, because they are all problematic individuals. Especially the Q_C is only 0.02 on database CB513.

4 Conclusion

In our work, we use the advanced PSSM as input data. The evolutionary information is important to improve the accuracy of secondary structure prediction. Three layer network are used as the prediction. After the first and the second NN ,the initial database are divided into two different database,then deal with them separately. Paying attention to the result of "good" and "bad", we can find that the PIs influence the classification certainty. When we deal with them separately, it's obvious that the Q_3 score improved.

By now, we decompose the initial database in the second NN, we will do experiment in the first NN and use better classified methods to divided initial database in the future.

To summarize, the paper's method can help to improve the prediction of secondary structure. Maybe our idea will become one of novel methods of breaking the current limitation in the field of secondary structure prediction.

Acknowledgment. This work was supported by the NSFC under grant No. 60573065, the Key Subject Research Foundation of Shandong Province and the Natural Science Foundation of Shandong Province under grant No. Y2007G33.

References

1. Kloczkowski, A., Ting, K.L., Jernigan, R.L., Garnier, J.: Combining the GORVAlgorithmWith Evolutionary Information for Protein Secondary Structure Prediction From Amino Acid Sequence [J]. PROTEINS: Structure, Function, and Genetics 49, 154–166 (2002)
2. Sun, X., Lu, Z.H., Xie, J.M.: The Basis of Bioinformation [M]. publication of TingHua, Beijing (2006)
3. Yuksektepe, F.U., Yilmaz, Q., Turkay, M.: Prediction of secondary structures of proteins using a two-stage method [J]. Computers and Chemical Engineering 32(2008), 78–88 (2008)
4. James, A.C., Geoffrey, J.B.: Evaluation and Improvement of Multiple Sequence Methods for Protein Secondary Structure Prediction [J]. PROTEINS: Structure, Function, and Genetics 34, 508–519 (1999)
5. Jones, D.T.: Protein secondary structure prediction based on position specific scoring matrices[J]. J. Mol. Biol. 292, 195–202 (1999)
6. Rost, B., Sander, C.: Prediction of protein secondary structure at better than 70% accuracy [J]. J. Mol. Biol. 232, 584–599 (1993)
7. Rost, B.: Review: Protein secondary structure prediction continues to rise [J]. Journal of Structural Biology 134, 204–218 (2001)
8. Lee, B.C., Kim1New, D.: Design of neural network Input and Output Vectors in the Protein Secondary Structure prediction [J]. Bioinformatics and Biosystems 1(4), 82–90 (2006)
9. Pagni, M., Cerutti, L., Bordoli, L.: An introduction to Patterns, Profiles, HMMs and PSI-PRED (2003)
10. Kabsch, W., Sander, C.: Dictionary of protein secondary structure: pattern recognition of hydrogen- bonded and geometrical features [J]. Biopolymers 22(12), 2577–2637 (1983)

Predicting Cytokines Based on Dipeptide and Length Feature

Wei He, Zhenran Jiang, and Zhibin Li

Department of Computer Science & Technology, East China Normal University, Shanghai 200241, China
lizb@cs.ecnu.edu.cn

Abstract. Cytokines are crucial intercellular regulators that have important physiological roles in a wide range of disease processes. The identification of new cytokines by computational methods can provide valuable clues in functional studies of uncharacterized proteins without performing extensive experiments. In this study, we developed a new prediction method for the cytokine family based on dipeptide composition and length distribution by using support vector machine (SVM). The cross-validation results demonstrated that cytokines could be correctly identified with an accuracy of 97% at family classification and 90% at subfamily recognition correctly, respectively. In comparison with existing methods in the literature, the present method displayed great competitiveness on identifying cytokines correctly.

Keywords: SVM, dipeptide composition, multi-class classification.

1 Introduction

Cytokines are cell intercellular messengers responsible for signaling an incredible variety of cell functions. They bind to specific cell receptors which then activate a signaling cascade and in the end a biological effect [1]. Cytokines can be divided into seven functionally distinct groups, including interleukins (IL), interferons (IFN), tumor necrosis factors (TNF), colony-stimulating factors (CSF), chemokines, and growth factors (GF) which include transforming growth factor beta (TGF-Beta), fibroblast growth factors (FGF), heparin binding growth factors (HBGF), and neuron growth factors (NGF) [2]. IL is the largest group of cytokines that stimulating immune cell proliferation and differentiation. This group includes IL-1 to IL-7, and in our work we only consider IL-6 as part of the dataset so as to compare the prediction performance of our method.

The characterization of biological functions of gene products is an important task in the post-genomic era. With the progress in genome sequencing of more and more model organisms, there exists an increasing gap between the amount of available sequence data and the experimental characterization of their functions. Therefore, it is necessary to develop automated methods for recognizing the new cytokines from primary sequence, which is a key task of bioinformatics in the post-genomic era. Although several algorithm, such as BLAST, HMM (Hidden Markov Model) and ANN (Artificial Neural Network), have been implemented to predict all kinds of

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 86 – 91, 2008.
© Springer-Verlag Berlin Heidelberg 2008

protein families, however, there are few literatures concerning the problem of predicting classes of cytokines based on their primary sequences information except the method developed by Huang [3] using dipeptide composition features.

In this paper, we developed a new SVM-based method for the prediction of cytokines based on dipeptide composition and the length feature of the sequences. The results showed that our method can get better performance for prediction of cytokines.

2 Materials and Methods

2.1 Datasets

The dataset used in this study, which is originally obtained from http://cytokine.medic.kumamoto-u.ac.jp/, consisting of 1173 cytokines belonging to the major classes listed on the web page. The eight classes are chemokine, FGF/HBGF, IL-6, LIF/OSM, MDK/PTN, NGF, TGF-Beta, and TNF. Then the highly homologous sequences were excluded from the dataset. Finally, 437 non-redundant sequences were obtained from the dataset, which includes seven classes except chemokine. To avoid a lack of adequate sequences, classes IL-6, LIF/OSM, MDK/PTN, and NGF are put into a single class.

2.2 Support Vector Machine

Support vector machines (SVMs) are a relatively recent method of machine learning based on structural risk minimization. The SVM builds a hyperplane separating positive examples and negative examples in higher-dimensional space [4]. This higher-dimensional space is called the feature space. SVMs learn the hyperplane that separates two classes with the maximum margin of separation. SVM uses kernel functions to map original data to a feature space of higher dimensions and locate an optimal separating hyperplane there. A kernel function of the dot product of the vectors is used to avoid representing the space explicitly. SVM training always seeks a global optimized solution and avoids over-fitting, so it has the ability to deal with a large number of features. This optimal hyperplane has been proven to guarantee the best performance. SVMs and related techniques have been greatly developed and applied to many bioinformatics areas [5].

The SVM was implemented using the freely software package LibSVM, which can be freely downloaded from http://www.csie.ntu.edu.tw/~cjlin/libsvm/ for academic use. The software enables the user to define a number of parameters as well as to customize inbuilt kernel functions. The experimentation was conducted using various types of kernels. All the kernel parameters were kept constant except for regulatory parameter C. The SVM was provided with fixed-length vector input. The fixed-length feature vectors were obtained from proteins of variable length using the coding method we developed.

2.3 Multi-class Classification

SVM is used based on two-class classification, but the classification of cytokine families is a multi-class classification problem. There are several algorithms for

multi-class classification such as 1-v-n, 1-v-1, and so on [6]. In this study, we select the method of 1-v-1, which constructs hyperplanes for every two classes. Assume that there are p classes, and thus $p(p-1)/2$ SVM are constructed. Each example is identified by the $p(p-1)/2$ SVM when training and testing, and then the class to which that example belongs is determined by voting mechanism. That is, each of the cytokines will achieve four scores, and it will be classified into the family that corresponds to the 1-v-1 SVM with the highest output score. Further, the numbers of some class of cytokine sequences are not large enough, which may cause the imbalance problem in the dataset. The 1-v-1 method will further aggravate the imbalance problem, while the 1-v-n method can improve the prediction accuracy for classes with fewer samples, and therefore the 1-v-n method was adopted finally [7-8].

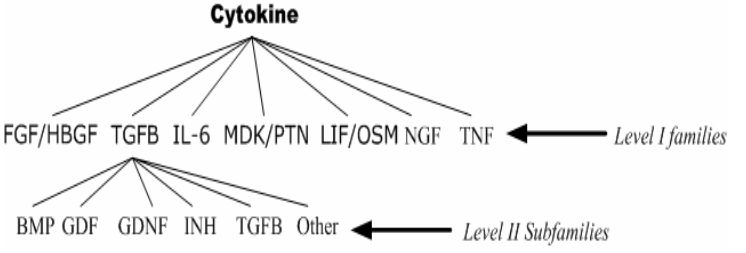


Fig. 1. Portion of the cytokine family tree showing the main classes of cytokine

2.4 Dipeptide Composition and Length Feature

The feature vectors are extracted by the means of dipeptide composition and the length from the sequences of cytokines. Dipeptide composition is a simplistic approach for producing patterns of fixed length from the protein sequences of varying length [3]. A number of literatures have demonstrated that dipeptide composition can provide global information on protein feature and give a fixed pattern length of 400. For instance, dipeptide composition has been widely used for predicting the subcellular localization of proteins and fold recognition effectively [9]. One of the important advantages is that dipeptide composition encapsulates information about the fraction of amino acids as well as their local order. Thus, the dipeptide composition of each protein was always calculated using equation as follows:

$$\text{Fraction of amino acid } i = \frac{\text{Total number of amino acid } i}{\text{Total number of amino acids in proteins}} \quad (1)$$

Where i can be any amino acid, "Total number of amino acid i" means in current cytokine sequence, the number of certain dipeptides of which the two amino acids are randomly taken from the 20 amino acids. The "Total number of amino acids in proteins" means always 400 that is 20×20 containing all possible compositions of dipeptides.

The length of a sequence is also important information for different classes of cytokines. In this study, the length of a sequence is divided by 10000 to form a dimension of input vectors. Thus, the final input vectors contain 401 dimensions.

2.5 Performance Evaluation

In this study, the performance of both sequence length and dipeptide composition based classifiers was evaluated through k-fold cross-validation. K-fold cross validation is an objective and rigorous testing procedure. The dataset of cytokines is randomly partitioned into k equal sets [12]. The training and testing of each SVM also called classifier was carried out k times using one distinct set for testing and other k-1 sets for training. The performance of SVM was evaluated by 10-fold cross validation in this paper. The results of five independent cross-validation experiments were averaged to get a fair evaluation. Four performance measures are used to assess the ability of the SVM classifier for the testing data, which are the overall classification accuracy (Acc), sensitivity (S_n), specificity (S_p) and Matthew's Correlation Coefficient (MCC) respectively defined as [10-11]:

$$\begin{aligned} Acc &= \frac{TP + TN}{TP + TN + FP + FN} \\ S_n &= TP \Big/ (TP + FN) \\ S_p &= TP \Big/ (TP + FP) \\ MCC &= \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}} \end{aligned} \tag{2}$$

where TP , TN , FP and FN are defined the numbers of true positive, true negative, false positive and false negative obtained from the prediction, respectively.

3 Results and Discussion

In order to evaluate the performance of our method, we compared it with that of the method developed by Huang using dipeptide composition features and evaluated with the same datasets. The performance of the method in distinguishing cytokine family is summarized in Table 1. The best results for each group are highlighted in bold face, which were achieved using RBF (Radial Basis Function) kernel with parameter gamma=20, punishment parameter C=1000 after being evaluated using a 10-fold cross-validation.

As seen from the Table 1, most of the prediction performances of our method are better than those of CTKPred. Table 1 also demonstrates that the average accuracy of our method was 2% higher as compared with dipeptide composition based classifier. The four threshold-dependent parameters of FGF/HBGF, Joint class (IL-6, LIF/OSM,

Table 1. The comparison results of cytokine family recognition with different method

	Dipeptide + length				Dipeptide			
	Sn	Sp	MCC	Acc	Sn	Sp	MCC	Acc
FGF/HBF	93.98 %	99.15 %	0.94	98.17 %	92.70 %	98.60 %	0.92	97.50 %
Joint class	97.06 %	99.73 %	0.97	99.31 %	91.00 %	99.70 %	0.94	98.40 %
TGF-Beta	96.84 %	97.56 %	0.94	97.25 %	97.40 %	94.70 %	0.92	95.80 %
TNF	95.79 %	97.95 %	0.93	97.48 %	94.00 %	98.80 %	0.94	97.70 %

Table 2. The comparison results of cytokine subfamily recognition with different method

	Dipeptide + length				Dipeptide			
	Sn	Sp	MCC	Acc	Sn	Sp	MCC	Acc
BMP	95.65 %	94.41%	0.87	94.71%	87.5 %	87.5 %	0.67	86%
GDF	83.87 %	97.47%	0.82	95.23%	82.4 %	82.4 %	0.76	93%
GDNF	87.50 %	97.11%	0.78	96.30%	75%	75%	0.86	98%
INH	89.29 %	100 %	0.94	98.41%	46.7 %	46.7 %	0.65	92%
TGF-Beta	65.71 %	96.10%	0.67	90.48%	100 %	100 %	0.96	99%
Other	75.76 %	94.23%	0.69	91.00%	66.7 %	66.7 %	0.56	84%

MDK/PTN, and NGF), and TGF-Beta were also superior to the corresponding ones of CTKPred accordingly.

The results of our cross-validation experiments recognizing cytokine in level 2 subfamilies are shown in Table 2. The above results obtained by our algorithm also indicate that cytokine proteins are considerably correlated with its dipeptide composition and the length properties, which can grasp most essential information reflecting cytokine identification.

In conclusion, we proposed an effective approach for predicting cytokine proteins using a supervised learning strategy. Cross-validated experiments show the prediction accuracies of cytokine family and subfamily recognition can reach 97%, and 90.9%, respectively. Future work of this study will concentrate on extracting more sequence features from primary sequence as feature vector of SVM. Therefore, it is possible to increase further the overall predictive accuracy by jointing other feature extracting methods and incorporate other informative features for predicting cytokine [13].

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grants No. 10771072). The authors appreciate the constructive criticism and valuables suggestions of the anonymous referees.

References

1. Luo, J.L., Tan, W., Ricono, J.M., Korchynskyi, O., Zhang, M., Gonias, S., Cheresh, D.A., Karin, M.: Nuclear Cytokine-activated IKKalpha Controls Prostate Cancer Metastasis by Repressing Maspin. Nature 446(7136), 690–694 (2007)
2. Schluns, K.S., Lefrançois, L.: Cytokine Control of Memory T-cell Development and Survival. Nat. Rev. Immunol. 3(4), 269–279 (2003)
3. Huang, N., Chen, H., Sun, Z.: CTKPred: An SVM-based Method for the Prediction and Classification of the Cytokine Superfamily. Protein Eng. Des. Sel. 18(8), 365–368 (2005)
4. Burges, C.: A Tutorial on Support Vector Machine for Pattern Recognition. Data Min. Knowl. Disc. 2, 121–167 (1998)
5. Pavlidis, P., Wapinski, I., Noble, W.S.: Support Vector Machine Classification on the Web. Bioinformatics 20, 586–587 (2004)
6. Karchin, R., Karplus, K., Haussler, D.: Classifying G-protein Coupled Receptors with Support Vector Machines. Bioinformatics 18, 147–159 (2002)
7. Lingras, P., Butz, C.J.: Reducing the Storage Requirements of 1-v-1 Support Vector Machine Multi-classifiers. RSFDGrC (2), 166–173 (2005)
8. Lingras, P., Butz, C.: Rough Set Based 1-v-1 and 1-v-r Approaches to Support Vector Machine Multi-classification. Inf. Sci. 177(18), 3782–3798 (2007)
9. Bhasin, M., Raghava, G.P.: Classification of Nuclear Receptors Based on Amino Acid Composition and Dipeptide Composition. J. Biol. Chem. 279, 23262–23266 (2004)
10. Hua, S., Sun, Z.: Support Vector Machine Approach for Protein Subcellular Localization Prediction. Bioinformatics 17, 721–728 (2001)
11. Huynen, M.A., Snel, B., von, M.C., Bork, P.: Function Prediction And Protein Networks. Curr. Opin. Cell Biol. 15, 191–198 (2003)
12. Zhang, S.W., Pan, Q., Zhang, H.C., Zhang, Y.L., Wang, H.Y.: Classification of Protein Quaternary Structure with Support Vector Machine. Bioinformatics 19, 2390–2396 (2003)
13. Xu, J.R., Zhang, J.X., Han, B.C., Liang, L., Ji, Z.L.: CytoSVM: An Advanced Server for Identification of Cytokine-receptor Interactions. Nucleic Acids Res. 35, W538–W542 (2007)

A New Unsupervised Approach to Face Recognition

Zizhu Fan and Ergen Liu

School of Natural Science, East China Jiaotong University, Nanchang 330013, China
zzfan3@yahoo.com.cn

Abstract. A new unsupervised approach to face recognition is proposed in this paper. Shape and color entropy is presented to descript face features. Firstly, images are pre-processed including face normalization and image segmentation and so on. Secondly, by using the information entropy theory, the method defines the color and shape entropy of the face images, respectively. Finally, an integrated similarity measurement framework is presented by computing mutual information between images according to these entropies. Compared with other methods of feature description, experiments indicate that this approach is more effective and efficient.

Keywords: mutual information; face recognition; image similarity; feature description.

1 Introduction

Face recognition is a research field that has attracted much attention in the past years due to its many applications in public surveillance, video annotation, multimedia and others. Face recognition systems usually fall into two categories [1]: verification and identification. Face verification compares a face image against a template face images, whose identity is being claimed. On the other hand, face identification compares a query face image against all image templates in a face database to determine the identity of the query face. Many techniques such as Principal Component Analysis (PCA) [2,3], Linear Discriminant Analysis (LDA) [4], Bayesian classifiers [5], Neural Networks [6], and Support Vector Machines (SVM) [7] have been proposed in this area. Although the existing algorithms are reported effectively, constructing new techniques for face recognition in practical applications is still an open and challenging problem [8].

In recent years, information entropy has been successfully applied in the field of image processing [9,10]. In this study, we use mutual information to match face images. This technique does not need training processing. So it can reduce computational cost. Face has many features. Shape is very important feature of face image. It can reflect local and some detail information of face. However, if it is used to depict image entire features, its effect is not good. It is possible to roundly describe whole image information by using shape character combining with color feature. By using color and shape features, this paper implements face recognition employing information entropy theory.

The organization of the paper is as follows: Section 2 describes traditional Shannon mutual information. It is improved in Section 3 that improves color mutual information

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 92 – 97, 2008.
© Springer-Verlag Berlin Heidelberg 2008

and presents shape mutual information. The measurement of face similarity is presented in Section 4 .Section 5 provides experimental results. Section 6 concludes with a discussion of the future work.

2 Shannon Mutual Information

In 1995, Viola [11] and Collignon [12] proposed respectively that mutual information was one of the image similarity measurements. It was very successful in application of calculation. Shannon mutual information is defined to be as the following.

Let H denote entropy and A be a face image representing image gray value. Suppose that the probability density function of image A is: $p_i = P(a = i)$, i =1 ,2 ,…, m

And Shannon entropy H(A) is defined as:

$$H(A) = -\sum_i p_i \log p_i \tag{1}$$

Then Shannon mutual information I(A,B) between image A and B is as following:

$$I(A,B) = \sum_{i,j} p_{ij} \log \frac{p_{ij}}{p_i p_j} = H(A) + H(B) - H(A,B) \tag{2}$$

where p_i and p_j are probability distributions of pixel values of A and B, respectively. p_{ij} is joint probability distribution of gray. H(A,B) is Shannon unite entropy of image A and B. In image match, Shannon mutual information depicts information amount of two images containing each other. If $I(A,B) = 0$, then image A is independent of image B. Similarity between two images becomes higher with the increasing $I(A,B)$.

3 Mutual Information of Color and Shape

In this paper, the traditional Shannon mutual information of image is ameliorated. In order to measure the similarity among images, mutual information of the color and shape among two or more images is computed firstly; then calculate the mean of these two types of mutual information. Thus, we focus on not only entire information of images, but local detail features. They effectively increase face recognition precision.

3.1 Improvement of Color Mutual Information

At present, there are lots of color-based (especially histogram-based) similarity measurements neglecting space information. So they don't get good effect. This paper gives an improved technology of color-base methods. That is, a face image is firstly segmented into some certain sub-areas before computing information entropy of image color. Secondly, whole mutual information between two images is calculated by using these sub-areas.

Suppose a gray image A can be segmented to n sub-areas: A_1, A_2… A_n .They satisfy:

$$A = \bigcup_{i=1}^{n} A_i, \qquad A_i \bigcap A_j = \phi \qquad 1 \le i, j \le n, i \ne j$$

Similarly, a gray image B can be segmented to m sub-areas: B_1, B_2, … B_m . They satisfy:

$$B = \bigcup_{j=1}^{m} B_j, \qquad B_i \bigcap B_j = \phi \qquad 1 \le i, j \le m, i \ne j$$

According to the formula (2), Shannon mutual information between two arbitrary sub-areas A_i, B_j which belong to image A and B respectively is to be

$$I(A_i, B_j) = H(A_i) + H(B_j) - H(A_i, B_j) \quad 1 \le i \le n, 1 \le j \le m$$

As for every A_i, it needs to compute $I(A_i,B_j)$ ($1 \le j \le m$). In order to effectively reflect color similarity among sub-areas, the biggest value of m similarities is selected. Thus, the improved mutual information $I_c(A,B)$ between image A and B is defined as:

$$I_c(A,B) = \sum_{i=1}^{n} \max_{1 \le j \le m} (I(A_i, B_j)) \ . \tag{3}$$

In general, as for a color image which has three color channels such as red, green and blue, we can define mutual information of these three color channels respectively. Then, the whole color mutual information between two images can be set to be the mean of these three types of mutual information.

3.2 Shape Mutual Information

The contour feature should be extracted so that shape mutual information between two images can be determined. In this paper, Canny operator is employed because it can nearly provide whole image edge.

Similar to color mutual information definition, the definition of shape entropy and shape mutual information can be obtained respectively through q_i and q_{ij} replacing p_i and p_{ij} in formula (1) and (2).So the shape entropy of image A can be define as[13]:

$$H_s(A) = -\sum_i q_i \log q_i \ . \tag{4}$$

Suppose a shape set S is obtained through edge detection from image A. And it contains k sub-shapes: $S_1,S_2,\ldots,S_k$. The shape set of image B is T which includes l sub-shapes: $T_1,T_2,\ldots,T_l$. Then Shannon mutual information $I(S_i,T_j)$ between two arbitrary sub-shapes S_i,T_j is defined as:

$$I_s(S_i, T_j) = H_s(S_i) + H_s(T_j) - H_s(S_i, T_j) \quad 1 \le i \le k, 1 \le j \le l$$

So, shape mutual information $I_s(A,B)$ between these two images is the following:

$$I_s(A,B)=\sum_{i=1}^{n}\max_{1\le j\le l}(I_s(S_i,T_j)) \tag{5}$$

4 Similarity Measurement and Face Match

From the Section 3, the entire mutual information $I^*(A,B)$ between two images is obtained:

$$I^*(A,B)=\alpha I_c(A,B)+\beta I_s(A,B) \tag{6}$$

where $0\le\alpha,\beta\le 1,\alpha+\beta=1$. We denote:

$$H_1=\alpha\sum_{i=1}^{n}H_c(A_i)+\beta\sum_{j=1}^{m}H_s(A_j)\quad,\ H_2=\alpha\sum_{i=1}^{k}H_c(B_i)+\beta\sum_{j=1}^{l}H_s(B_j)\ .$$

where $H_c(A_i)$ and $H_c(B_i)$ are color entropies of sub-images A_i and B_i in images A and B, respectively; $H_s(A_j)$ and $H_s(B_j)$ are shape entropies of sub-images A_j and B_j, respectively. Thus, the similarity $S(A,B)$ between two images is defined as the following:

$$S(A,B)=\frac{I^*(A,B)}{\max(H_1,H_2)} \tag{7}$$

α,β are similar to the values of formula (6). We can prove: $0\le S(A,B)\le 1$. In order to match face images, it is necessary to experientially set a threshold. If similarity between two face images is bigger than this threshold, we can conclude that these two face images belong to one person.

5 Experimental Results

In this section, for the purpose of simple calculation and convenient comparison, we elaborate on the experimental findings for two well-known and commonly used grayscale face databases as shown in Fig.1. We contrast the results of the mutual information approach with the previous techniques. The ORL [14] database comprises 400 face images coming from 40 individuals; the pictures were taken in different environments. The total number of images for each individual is equal to 10. These images vary in position, rotation, scale, and facial expression. For some individuals, the images were taken at different times, varying facial details. Each image was digitized and stored as a 112×92 pixel array whose gray levels ranged between 0 and 255. The Yale Face Database [15] contains 165 grayscale images in GIF format of 15

(a) (b)

Fig. 1. Two typical face image databases: pictures in (a) are from ORL face image database; the right images displayed in (b) are chosen from Yale face image database

Table 1. Comparison of mean for recognition rates

	Eigenface (PCA)	Fishface (PCA+LDA)	Fuzzy fisherface (Fuzzy+PCA+LDA)	Mutual information (Color + shape)
ORL	0.939	0.950	0.960	0.972
Yale	0.762	0.903	0.930	0.945

individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink.

In the experiment, face images were segmented using the region growing algorithm in advance to extract the color feature. In the process of computing image color entropy, we focused on big regions. If gray variance of one region is less than that of whole image, then its color entropy will not be computed, because it can not enhance recognition rate. Moreover, the calculation can increase computational cost. Like the color, a small shape can also be neglected. In the experiment, if a color block or shape block has one percent of image area, it was overlooked when we computed mutual information. Then detected their contours employing Canny arithmetic operators to obtain image shapes. And finally, mutual information among images was calculated.

In the formula (6), the color and the shape have same weight. Both of them were 0.5. Face image matching threshold was experientially set to be 0.7. This threshold also can be set through computing similarity among some images chosen in advance. So our method almost does not need any training set. We compared our method with other three techniques mentioned in literature [16]: PCA, PCA+LDA, Fuzzy+ PCA+ LDA. Literature [16] discussed three cases. Table 1 gave mean rates for these cases using two typical face databases. From Table 1, effective of experiment on ORL is better than that of Yale database because Yale face images are more complicated. Among four approaches, Eigenface got the worst result; our method is best. The rates of recognition on ORL and Yale databases are 0.972 and 0.945, respectively.

6 Conclusion

In this paper, a new unsupervised face recognition approach based-on mutual information is proposed. We extended the concept of image entropy and gave two definitions: the color and shape entropy of image. Compared with other supervised

methods, our algorithm can be implemented without a training set. It only needs an experiential threshold. Experiments show it is promising. In order to increase effective and efficiency of face recognition, our future work will focus on local features and texture information of face images.

Acknowledgements. The authors would like to thank Natural Science Foundation of Jiangxi Province of China (No. 0611009) and the education department of Jiangxi province for financially supporting this research under Contracts No. GJJ08254.

References

1. Andrea, F.A., Michele, N., Daniel, R., Gabriele, S.: 2D and 3D face recognition: A survey. Pattern Recognition Letters 28, 1885–1906 (2007)
2. Turk, M.A., Pentland, A.P.: Eigenfaces for recognition. J. Cognit. Neurosci. 3(1), 71–96 (1991)
3. Moghaddam, B.: Principal manifolds and probabilistic subspaces for visual recognition. IEEE Trans. Pattern Anal. Machine Intell. 24(6), 780–788 (2002)
4. Belhumeur, P.N., Hespanha, J., Kriegman, D.J.: Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell. 19(7), 711–720 (1997)
5. Liu, C., Wechsler, H.: A unified Bayesian framework for face recognition. In: Proc. Internat. Conf. on Image Processing (ICIP 1998), pp. 151–155 (1998)
6. Sung, K.K., Poggio, T.: Example-based learning for view-based human face detection. IEEE Trans. Pattern Anal. Machine Intell. 20(1), 39–51 (1998)
7. Tefas, A., Kotropoulos, C., Pitas, I.: Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication. IEEE Trans. Pattern Anal. Machine Intell. 23(7), 735–746 (2001)
8. Johnny, K.C., Ng., Z.Y.Z., Yang, S.Q.: A comparative study of Minimax Probability Machine-based approaches for face recognition. Pattern Recognition Letters 28, 1995–2002 (2007)
9. Lu, X.S., Zhang, S., Su, H., Chen, Y.Z.: Mutual information-based multimodal image registration using a novel joint histogram estimation. Computerized Medical Imaging and Graphics 32(3), 202–209 (2008)
10. Suyash, P.A., Tolga, T., Norman, F., Ross, T.W.: Adaptive Markov modeling for mutual-information-based, unsupervised MRI brain-tissue classification. Medical Image Analysis 10(5), 726–739 (2006)
11. Viola, P., Wells, W.: Alignment by maximization of mutual information. In: Proceedings of the 5th International Conference on Computer Vision, Boston, MA, pp. 16–23 (1995)
12. Collignon, A., Maes, F., Vandermeulen, D., et al.: Automated multimodality image registration using information theory. In: Proceedings of the Information Processing in Medical Imaging Conference, Dordrecht, pp. 263–274 (1995)
13. Fan, Z.Z., Zhou, S.C.: Image Retrieval Based on Shape Entropy. Journal of Computer Application & Research 24(9), 309–311 (2007) (in Chinese)
14. ORL face database (2008), `http://www.uk.research.att.com/facedatabase.html`
15. Yale face database (2007), `http://cvc.yale.edu/projects/yalefaces/yalefaces.html`
16. Kwaka, K.C., Witold, P.: Face Recognition Using a Fuzzy Fisherface Classifier. Pattern Recognition 38, 1717–1732 (2005)

Human Motion Analysis Using Eroded and Restored Skeletons

Rong Hu, Chucai Yi, and Hongyuan Wang

Electronics and Information Department, Huazhong University of Science and Technology, Wuhan, Hubei Province, China
{hr,chucai2000}@smail.hust.edu.cn,
wythywl@public.wh.hb.cn

Abstract. Skeleton has been found a very useful feature in human motion analysis and gait recognition. However, it is usually hard to obtain the right feature skeleton because of shape variance of human silhouette and noise. Due to those reasons, unexpected skeleton branches will exist together with the main components of human body. To delete the unexpected branches, a silhouette reconstruction method is proposed to cope with the shadow problem in foot area, while a skeleton pruning technique using skeleton erosion and restoration is proposed to extract the main body parts. The final skeleton can be used as a high level gait feature to carry out further motion type analysis.

Keywords: Skeleton, Erosion, Restoration, Motion Analysis, Gait.

1 Introduction

Early research has pointed out that the motion of an object can be analyzed by examining its cycles in movement [11], and the object's information such as age, weight and gender can be reflected by its stride parameters [7]. In human motion analysis applications, the periodic variations of parts of body are usually desired to be obtained. As a higher level of feature than silhouette, skeleton is able to describe the temporal posture of human more compactly and efficiently. It can also be used as the model to represent the topological structure and shape of persons, like other model based methods [1], [2], [3] in human gait recognition applications.

However, it is hard to get the desired feature skeleton which contains only the expected parts due to the complexity of silhouette, such as clothing, shadow, and noise. Further more, skeleton extracting methods [4], [5] have strict requirement for silhouettes. Shadows and noise will bring in unexpected influences. These problems should be settled before motion analysis begins.

First of all, the unexpected parts have to be removed from silhouette such as shadows and moving objects in background which may change the topological structure of human skeleton. Then skeleton of silhouette is extracted using the extracting method in [4], which is followed by a skeleton pruning procedure. The rest of sections are organized as follows: Section 2 exhibits the human silhouette reconstruction procedure; Section 3 explains the skeleton erosion and restoration methodology; Section 4 gives the analysis on the experiment result and our future work.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 98 – 105, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2 Silhouette Reconstruction

We use the simple background subtraction scheme as proposed in [6] to get the rough silhouettes. However, silhouettes extracted by this method are not immediately suited for skeletonization if they are not recorded in laboratorial environment. Let a library be a collection of manual silhouettes consists of different motion types, and there is a sequence of manual silhouettes for each motion type consisting one motion period. Suppose the library is big enough to contain most of the motion types, and for each motion type, there are enough stances (phases) to represent the motion variance in a period. Let a probe sequence be the sequence of silhouettes to be reconstructed. For each probe sequence, the number of stances in one period should be fewer than that in the library sequence.

Through a training step, the most similar motion types are found from the library and the probe silhouettes are aligned with them in phase. The motion period is calculated by the method in [6] for both the probe sequences and the library sequences. Each period starts from back foot leaving land and ends with front foot touching land. Figure 7 just illustrates half of a gait period. Since there are more stances in library sequence, the library sequence is partitioned into N subsets, while N is the number of stances in probe sequence, and each probe silhouette corresponds to one subset, figure 1(a). The numbers of stances in each subset are equal because two consecutive subsets may have the same stance.

Let the similarity of two silhouettes be defined as: $S_{A,B} = (A \cap B)/(A \cup B)$, here A and B stands for pixel set of silhouette. The library subsets are rearranged by calculating the similarity between each silhouette in library subset and the probe silhouette corresponding to this subset or the neighboring subsets. For each library silhouette, if its similarity with probe silhouette in the neighboring phases is greater than the similarity with current phase, the library silhouette is moved into neighboring subset, else wise, it remains in the current subset. Rearrange result is illustrated in figure 1(b).

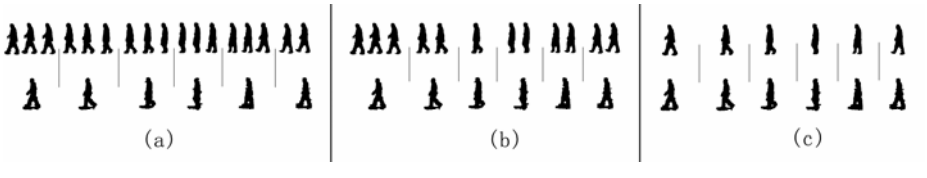


Fig. 1. The upper row are silhouettes from library, the lower row are silhouettes from probe. (a): roughly aligned in phase; (b): rearranged sequence; (c): most similar silhouette sequence.

At last, the most similar silhouette in each subset is picked up, figure 1(c). After picking up corresponding silhouette for each silhouette in probe sequence, the average similarity of two sequences can be calculated, and it can be used as the criterion to choose the most similar motion types. For reconstruction purpose, just one library sequence is not enough because it can not be exactly the same as the probe one. So the average information is needed to include every possibility in a gait sequence. For

each phase, all the selected library silhouettes are summed up to form an added-average silhouette.

Let set A be the probe silhouette and set B be the template for A to be reconstructed to. There comes the following definition:

1. $A \bigcap B$ is the common information which is useful;
2. $A \bigcap (A \bigcup B - A \bigcap B)$ is the unexpected information to be deleted;
3. $B \bigcap (A \bigcup B - A \bigcap B)$ is the missing information to be added in.

According to these definitions, we get three images, as shown in figure 2(b), (c) and (d). Note that they are all gray level because template is an average image of silhouettes. In figure 2(c), the higher the gray level is, the more possible the corresponding pixel in probe will transit to background. The shadow is usually the highest part in this image because none of the library sequences have shadow in the foot area even if we use the average image of library silhouettes. In the same way, the higher the gray level in figure 2(d) is, the more possible the corresponding pixel in library will transit to foreground.

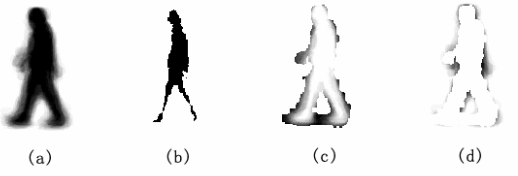


Fig. 2. Image obtained from definition 1, 2 and 3. (a): added-average silhouette; (b): useful information; (c): unexpected information; (d): lost information

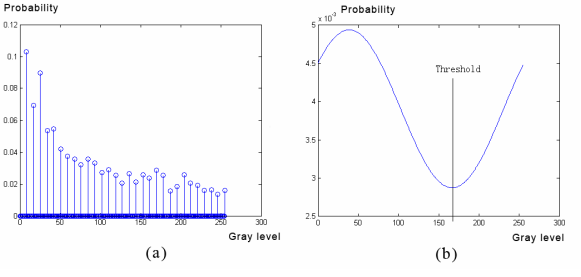


Fig. 3. (a): Histogram of gray level image; (b): filtered histogram for threshold

The thresholds for these transitions are determined as followed. First we get the histograms of the gray images, figure 3(a). Then filter them out using a low-pass filter with a low cut-off frequency. The larger the frequency is, the more valley/peaks would be. Normally we choose a small value so that only several valley/peaks exist. The gray level of the last valley is selected as the threshold for reconstruction. The right part of threshold in figure 3(b) represents the clusters of points which are of high

probability to exist or not, so they are transformed. The final result of reconstruction from the thresholds is shown in figure 7.

3 Skeleton Erosion and Restoration

Many efforts have been made in skeleton pruning [9], [10]. Each has its own drawbacks or limitations, such as the topology of the global shapes must be known before skeletonization for the fixed topology skeleton and the position of the obtained skeleton is not accurate. Our pruning method only focuses on the lengths of the branches in skeleton and tries to remove the unnecessary skeleton branches whose lengths are less than the main components. Firstly, some definitions from [8] is introduced: The endpoint is the point which has only one neighbouring point in the skeleton; the connection point is the point which has two neighbouring points in the skeleton; and the branch point is the point which has more than two neighbouring points in the skeleton. Here, we extend the definition into the discrete image cases.

Let N(p) be the set of neighbors of point p and it is termed the neighborhood of p, see figure 4(a). p[0], p[2], p[4], p[6] are north(p), east(p), south(p), west(p) respectively and p[1], p[3], p[5], p[7] are north-east(p), south-east(p), south-west(p), north-west(p) respectively. The value for p[k] is either zero or one, and the weight number for p is denoted by:

$$WN(p) = \sum_{k=0}^{7} p[k] * 2^k$$

For a 3*3 window, the weight number ranges from 0 to 255, each weight number denotes a special neighborhood distribution of p, and a weight number set denotes a set of neighborhood distributions of p.

Definition 1. The weight number set ES = {1,2,3,4,6,8,12,14,16,24,32,48,56,64,96, 128,129,131,192,224}. An endpoint is the point whose weight number is in ES.

Definition 2. The weight number set of connection point is CS. A connection point is the point whose weight number is in CS. Here the weight numbers in CS are omitted. The inferring process is explained as follows.

We can see from figure 4(b) and (c), there should be at least one point in the neighborhood of an endpoint p. If this point is in the odd neighborhood, such as p[1], p[3], p[5], p[7], no other point should exist in the neighborhood of p any more. If this point is in the even neighborhood, such as p[0], p[2], p[4], p[6], additional points can exist in the neighborhood of p which are 4-connected with this point.

We can see from figure 4(d), (e) and (f), there should be at least two points in the neighborhood of a connection point p and these two points are not 4-connected. If any of the points is in the even neighborhood, such as p[0], p[2], p[4], p[6], then additional points can exist in the neighborhood of p which are 4-connected with this point, otherwise, no more neighbors should exist around p. Note that no loop is allowed as shown in Figure 4(g), which only illustrates one of the four cases.

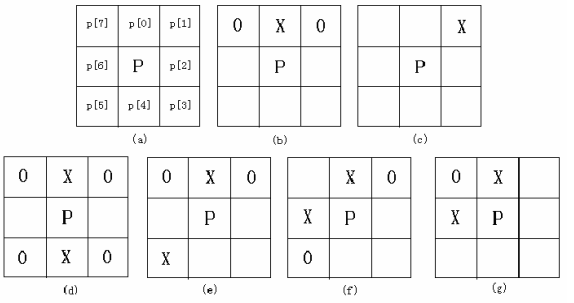


Fig. 4. X represents a fixed point, and O represents an optional point. No points exist in blank grids.

3.1 Skeleton Erosion

The operation for skeleton erosion is deleting all the endpoints in the skeleton iteratively until the terminating condition occurs.

Step 1: calculate the weight number for each point in the skeleton;
Step 2: find the endpoints by definition 1 and delete them;
Step 3: update the weight numbers for the neighbors of deleted endpoints;
Step 4: terminating condition occurs? Yes, go to step 5. No, go to step 2;
Step 5: start restoration.

We have to introduce a new term named erosion degree which is the iteration times for erosion and the terminating condition. A branch is a trace from an endpoint to another endpoint or branch point. Any branch whose length is less than the erosion degree will vanish, figure 5(a). A critical situation happens when loops exist in the skeleton. If there is no loop, and the erosion degree is big enough, the whole skeleton will disappear. Otherwise, the loop will exist still, as in Figure 5(b). To eliminate the loop, we first find the remaining points with a big erosion degree and mark them, and then does the erosion step again. If any neighbor of the endpoint to be deleted is marked, the neighbor of the endpoint point is also deleted, and the erosion goes on, figure 5(c).

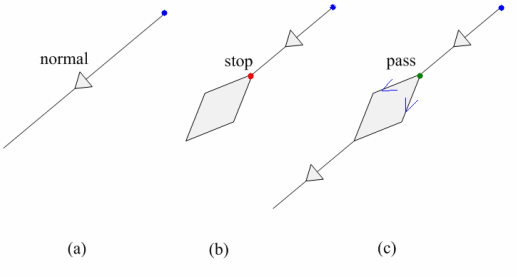


Fig. 5. (a): No loop; (b): has a loop, erosion stops; (c): break the loop, erosion goes on

3.2 Skeleton Restoration

For each iteration step in section 3.1, we record the deleted endpoints DE(k). DE(k) is the set of deleted endpoints, k is the index for erosion. For example: DE(1), DE(2), …, DE(k). The restoration steps are as followed:

Step 1: k is set to the erosion degree;
Step 2: for each point in DE(k), find those points whose relationships with the existing skeleton end points satisfy definition 1;
Step 3: form a new skeleton as the union of the old skeleton and the found endpoints;
Step 4: k is decreased by one. Is k equal to zero? Yes, end of loop. No, go to step 2.

The skeleton pruning result is shown in figure 6. The left part in this figure is the raw skeleton of a horse and a leaf; the right part is the pruning result.

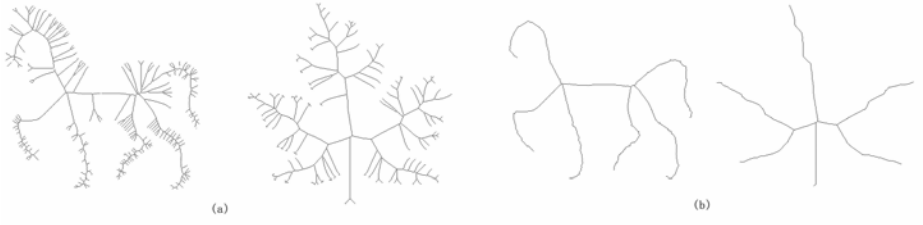


Fig. 6. (a): Original skeleton; (b): eroded and restored skeleton

4 Experiment Analysis

The upper row of figure 7 is raw silhouettes by the simple background subtraction method. Note the shadow in feet area, which will cause a significant change in feet skeleton. Sometimes the shadow even connects the feet skeleton, and it means the topological structure of legs changes completely. Noise on the boundary of silhouette will also introduce many unexpected branches. The bottom row of figure 7 is the reconstructed silhouettes. We can see that shadows and significant noise are removed from silhouettes.

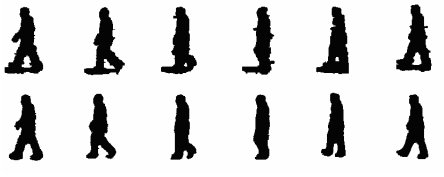


Fig. 7. The upper row is original silhouettes; the lower row is corresponding reconstructed silhouettes

In silhouette reconstruction step, we suppose that the most similar silhouettes selected from library include very possible temporal gesture of the probe silhouette for each phase in the period. However, if the probe silhouette is so special that some part

of the silhouette, which is different from all the selected library silhouettes, will be deleted. It is usually the case for shadow or other significant noise as we expect. In the histogram of figure 3, the horizontal axis stands for the gray level which can also be explained as the probability of state change from foreground to background for each pixel in probe silhouette image. Here, the probability ranges from 0 to 255. The vertical axis represents the number of occurrence for each probability of points. The cutoff frequency for determining gray-level threshold is selected according to how much information you want to delete from the image. Normally, the larger the cutoff frequency is, the less the points to be deleted would be.

Thinning method is used to exact the human skeletons. Since this method is sensitive to noise, unexpected branches exist even if the silhouette has been reconstructed. Note from figure 8(a), the unexpected branches are shorter than the main body part. The longest trace in skeleton is the one from head to feet, and this is what we need for human motion analysis. The skeleton pruning result of human gait is shown in figure 8(b), from which we can see that hand/arm of human do not exist sometimes. This is because hand/arm area is often mixed up with body area in silhouette, so the obtained skeleton part for this area is short or even does not exist. It is hard to distinguish the hand/arm skeleton branch from other branches so the hand/arm skeleton branch is deleted too. In our experiment, erosion degree is set to half of the width of silhouette image. Basically, there is only one new endpoint connecting an old endpoint in the process of skeleton restoration. However, it is still possible that two or more new endpoints exist, because several branches disappear at the same time and at the same endpoint in the process of skeleton erosion, which takes a really low possibility. In this case, one eroded trace may become two traces or more after its restoration. The worst situation happens when some unexpected branch by noise is longer than the main body part. Error occurs when we remove the desired branch while keeping the unexpected branch in the skeleton. This problem should be solved in our future work.

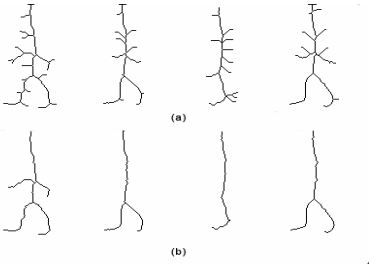


Fig. 8. (a): Human skeletons with unexpected branches; (b): reconstructed skeletons

After the skeleton is cleaned up, main components such as legs and feet become clear, and the distances between feet can be collected. Figure 9 illustrates the feet distance variations of four people for two periods. Conclusions like stride length, cadence and asymmetry in the gait period can be drawn from the cyclic variation of this distance, which will tell the differences between people, like age and gender [7]. More matching and classifying methods using skeleton are needed in the future work.

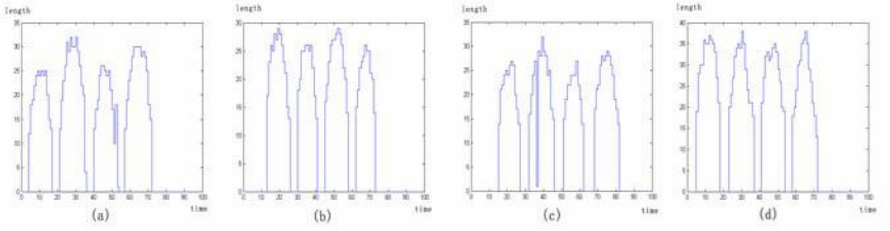


Fig. 9. Feet distance variation of four people for two periods

References

1. Wagg, D.K., Nixon, M.S.: On Automated Model-based Extraction and Analysis of Gait. In: Proceedings - Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 11–16 (2004)
2. Johnson, A., Bobick, A.: A Multi-view Method for Gait Recognition Using Static Body Parameters. In: Third International Conference on Audio- and Video-Based Biometric Person Authentication, pp. 301–311 (2001)
3. Cunado, D., Nixon, M.S., Carter, J.N.: Automatic Extraction and Description of Human Gait Models for Recognition Purposes. Computer Vision and Image Understanding 90(1), 1–41 (2003)
4. Gilles, B.: Parallel Thinning Algorithm for Medial Surfaces. Pattern Recognition Letters 16(9), 979–986 (1995)
5. Hassouna, M., Sabry, F., Aly, A.: Robust Centerline Extraction Framework using Level Sets. In: Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 458–465 (2005)
6. Phillips, P.J., Sarkar, S., Robledo, I., Grother, P., Bowyer, K.: The Gait Identification Challenge Problem: Data Sets and Baseline Algorithm. In: Proceedings - International Conference on Pattern Recognition, vol. 16(1), pp. 385–388 (2002)
7. BenAbdelkader, C., Cutler, R., Davis, L.: Person Identification Using Automatic Height and Stride Estimation. In: Proceedings - International Conference on Pattern Recognition, vol. 16(4), pp. 377–380 (2002)
8. Liu, W.Y., Liu, J.T.: Objects Similarity Measure Based on Skeleton Tree Descriptor Matching. Journal of Infrared and Millimeter Waves 24(6), 432–436 (2005)
9. Bai, X., Latecki, L.J., Liu, W.Y.: Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(3), 449–462 (2007)
10. Golland, P., Grimson, E., Eric, L.: Fixed Topology Skeletons. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 10–17 (2000)
11. Tsai, P., Shah, M., Ketter, K., Kasparis, T.: Cyclic Motion Detection for Motion Based Recognition. Pattern Recognition 27(12), 1591–1603 (1994)

Activities Prediction of Drug Molecules by Using the Optimal Ensemble Based on Uniform Design

Yue Liu, Yafeng Yin, Zaixia Teng, Qi Wu, and Guozheng Li*

School of Computer Engineering & Science, Shanghai University,
Shanghai, 200072, China
gzli@shu.edu.cn

Abstract. Neural network ensemble is a powerful tool for simulating the quantitative structure activity relationship in drug discovery because of its high generalization ability. However, the architecture of the ensemble and the training parameters of individual neural networks are closely relative to the generalization performance of the ensemble and the convenience of the creation of the ensemble. This paper proposes a novel creation algorithm for neural network ensemble, which employs uniform design to guide users to design the ensemble architecture and adjust the training parameters of individual neural networks. In addition, this algorithm is applied to produce neural network ensemble for predicting activities of drug molecules, which is a convenient way to achieve better results than commonly used bagging methods.

Keywords: Neural Network Ensemble, Uniform Design, Drug Design.

1 Introduction

Quantitative Structure Activity Relationship (QSAR) models help to instruct new drugs design. The prediction of new drug molecules using computational technology will economize quantity experiment costs and improve efficiency of drug design [1]. The common applied methods in QSAR analysis include multiple regression, k-neighbors, partial least square regression, decision trees, Kriging method, artificial neural networks, semi-supervised learning method like co-training, and so on. In recent years, SVM (Support Vector Machine) and ensemble learning methods have been used to model this problem and obtained satisfactory results [2].

Neural network ensemble is a learning paradigm, in which a collection of a finite number of neural networks is trained for the same task [3]. It has become a hot topic in the machine learning community because of its high generalization ability. Recently, various ensemble methods have been proposed. However, the architecture of the ensemble (the number of individual neural networks, the number of hidden nodes of individual neural networks) and the parameters for training the individual neural networks are all determined manually and fixed during the training process. As well known, the ensemble architecture design involves a tedious trial-and-error process for the real world problems and depends on the users' prior knowledge and experience

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 106 – 113, 2008.
© Springer-Verlag Berlin Heidelberg 2008

[4]. In order to apply the neural network ensemble to QSAR successfully, it is crucial to develop new ensemble methods, which can be conveniently employed to construct the ensemble's architecture and easily adjust the training parameters of the individual neural networks. As far as we know, Yao's group proposed a method called CNNE (Constructive Neural Network Ensemble) [4] for feed forward neural networks based on negative correlation learning, in which both the number of neural networks and that of hidden nodes in individual neural networks are determined using a constructive algorithm. We also proposed a method named CERNN (Constructive Ensemble of RBF Neural Networks) [5] in which a constructive algorithm and a nearest-neighbor clustering algorithm are utilized to make the users without any experience of neural networks expediently and automatically construct the ensemble. Akhand et al. [6] devoted themselves to the research on the constructive ensemble. Zhou et al. [7] and Li et al. [8] employed a genetic algorithm and clustering algorithm individually to determine the number of individual neural networks of the ensemble, these methods are called selective ensemble.

Uniform design [9] is a new experimental method which is applied to arrange experiments scientifically, reduce the time-cost of experiments with good results. This paper proposes a novel algorithm named ORNNEUD (Optimal RBF Neural Network Ensemble based on Uniform Design), which employs uniform design to guide users with little experience of neural networks to design the ensemble architecture and adjust the training parameters of individual neural networks. At the same time, the nearest-neighbor clustering algorithm is used to create the heterogeneous individuals with different hidden nodes. The appropriative individuals are selected to construct the final neural network ensemble. At last, ORNNEUD is applied for modeling of QSAR, and gives better results comparing with the previous proposed selective ensemble learning algorithms.

The rest of this paper is organized as follows. In Section 2, details of the proposed novel algorithm are described. Experiments are performed on a drug molecular activity dataset in Section 3. Section 4 concludes.

2 An Optimal RBF Neural Network Ensemble Based on Uniform Design

The main ideas of the proposed ORNNEUD (Optimal RBF Neural Network Ensemble based on Uniform Design) are briefly summarized as follows:

(1) Get different training samples for each network by using the Bootstrap resampling method, which enlarges the diversity among individual networks.
(2) The nearest-neighbor clustering algorithm is introduced to determine the neuron number in the hidden layer of RBFNN by using its individual samples, which constructs the individual network structure and further increases the diversity.
(3) Use uniform design to determine the parameters of network, such as learning rate, maximum training epoch, gauss width, etc. for the purpose of reducing the network's generation error.
(4) With the guidance of results of uniform design, we can select the most proper RBFNN to compose the ensemble.

Details of ORNNEUD can be described as the following steps:

Step 1: Carry out multi-factor and multi-level uniform design experiments, get t groups of different parameters combination for the training process of each individual RBFNN, and construct the uniform design table.

Step 1.1: Confirm the factors in uniform design table. In this paper, network learning rate, maximum training epoch, and gauss width are chosen as factor, while the correct classification rate $(1-\mathrm{E})$ is selected as evaluating index, where E represents the error classification rate that is defined as formula (1).

$$E = \frac{1}{N}\sum_{n=1}^{N} C(n), \text{if } O(n) = D(n) \text{ then } C(n) = 0 \text{ else } C(n) = 1. \tag{1}$$

where $\{\mathrm{D}(n) | n = 1,..., N\}$ denotes the teacher information of the test data set, $\{O(n) | n = 1,..., N\}$ represents the output of ensemble, N is the number of samples in the test data set.

Step 1.2: Choose different levels for each factor by experience and construct the factor-level table.

Step 1.3: Select the best uniform design table and then get t different combination of parameters called schemes for training RBFNN.

Step 2: Train and test t different RBFNNs based on the t schemes. The detail process to acquire one RBFNN can be described as follows:

Step 2.1: Generate one training set by using the bootstrap method.

Step 2.2: Use the nearest neighbor clustering algorithm to determine the number of neurons of the hidden layer for RBFNN [5] and then train the RBFNN.

Step 2.3: Use the test set to test performance of RBFNN.

Step 3: Analyze the results of step 2 and delete the RBFNN whose classification correction rate does not meet our requirement under the condition of setting part parameters. If any RBFNN is removed, go to step 2 to build a new one as substitution.

Step 4: Use the multiple linear regression method to analyze t RBFNNs generated in the previous steps and find the scheme which performs the best.

Step 5: Get new RBFNNs by using their corresponding best scheme.

Step 6: Repeat steps 2-5 to generate as much individual RBFNNs as we wanted (in this paper, 50 is required).

Step7: Integrate the RBFNNs whose performances are enhanced after optimization to construct the final ensemble model.

3 Comparative Studies on a Drug Molecular Data Set

3.1 Data Preprocessing

The Mutagenicity (available on http://www.niss.org) data set is used. Totally 1835 samples are included in this data set and there are 618 structure parameters including 47 CONS descriptors, 260 Topological Indexes, 64 BUCT descriptors and 247 FRAG descriptors. Through pre-analysis of this data set, we select 10 descriptors as input and the activity label as output.

3.2 Experimental Setup

In this paper, 75% samples are used as the training set, while the other 25% as the test set. Network learning rate, training epoch and gauss width of individual neural networks in the ensemble are chosen as uniform design factors. Here, each factor has four levels. The uniform design table is set up as Table 1 and the corresponding experiment schemes are built as Table 2.

Table 1. Factor-level combination of drug molecular activity analysis

Factor / Level	*A* Training epoch	*B* Learning rate	*C* Gauss width
1	1000	0.1	0.7
2	2000	0.3	0.8
3	3000	0.5	0.9
4	4000	0.9	1.0

In table 2, Parameters 1, 2, 3, 4, and 5 are used to denote the different combination of parameters for training RBFNN in the experiments. For example, Parameter 1 means that Training epoch=3000, Learning rate=0.5, and Gauss width=1.0.

Table 2. The uniform design table of drug molecular activity analysis

Factor / Level	A Training epoch	B Learning rate	C Gauss width	Evaluating Index Prediction accuracy (%)
Parameters 1	3000	0.5	1.0	
Parameters 2	2000	0.9	0.7	
Parameters 3	1000	0.1	0.9	
Parameters 4	4000	0.3	0.8	
Parameters 5	2000	0.5	0.9	

3.3 Experimental Results and Analysis

To validate the effectiveness of the novel method ORNNEUD, it is compared with other five methods. Details of these algorithms are described as follows.

- **RBFNN & ORNNU**D (Optimal single RBF Neural Network based on Uniform Design). The original data set is used directly to train and test one RBF neural network with parameter combination of maximum training epoch 1000, network learning rate 0.3 and gauss width 0.7. An under-fitting RBFNN is trained. ORNNUD, in which the uniform design method is employed to select the best parameter combination in Table 2, is performed. As shown in Table 3, both RBFNN and ORNNUD work badly. Their prediction accuracies are only 49.89%.
- **Bagging.** 50 different sample subsets are generated by bootstrap method and each subset is used to build one RBFNN with the same training parameters. For

eliminating the effect on ensemble precision caused by diverse parameters combination, we use five parameter combinations shown in Table 2. The experimental results are 60.22% (Parameters 1), 63.66% (Parameters 2), 61.29% (Parameters 3), 60.22% (Parameters 4), and 61.08% (Parameters 5) respectively. It indicates that bagging makes a progress in prediction accuracy under different parameters than RBFNN and OENNUD. It is also observed from Fig.1 that 27 (i.e. the No. 3, 4, 5, 6, 8, 10, 14, 16, 17, 21, 22, 24, 28, 29, 30, 31, 34, 36, 37, 40, 41, 43, 45, 46, 48, 49, and 50) individual RBFNNs have poor prediction accuracy on test data set which is 49.89%. We call these individuals "bad individuals".

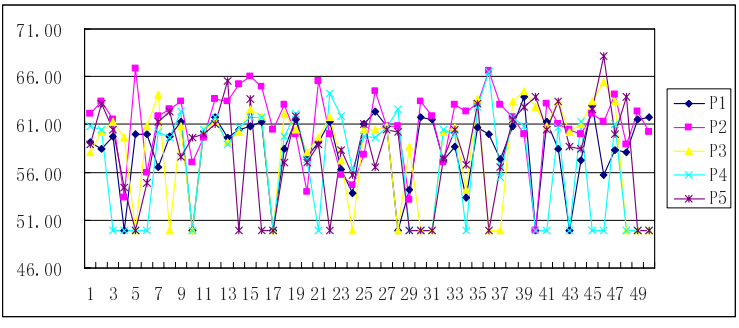


Fig. 1. Predictive accuracy of individual RBFNNs of bagging on the test data set (%)

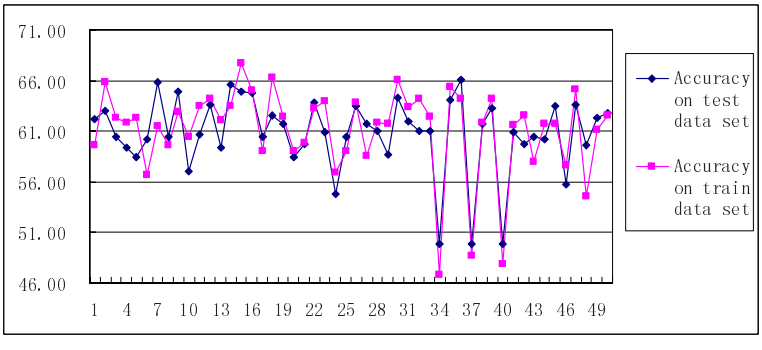


Fig. 2. Predictive accuracy of individual RBFNNs of OBUD (%)

- **OBUD** (Optimal Bagging ensemble based on Uniform Design). Uniform design method is used to optimize the individual RBFNNs of bagging, and then integrate them to construct the new ensemble named OBUD. Here the training parameters of each individual RBFNN are the same as those in bagging. From Fig.2, we can see that only the 34th, 37th, and 40th individual RBFNNs are under-fitting.
- **ORNNEIUD** (the Optimal RBF Neural Network Ensemble with Individuals' selection based on Uniform Design). There are 27 "bad" individuals in bagging. We remove the "bad" ones and generate new ones as substitution using the same way till the precision of 50 new individuals are larger than 50%. And the new ones are

named as “good individuals”. The performances of the “good” individuals are shown in Fig.3. Then, these “good” ones are optimized by using uniform design and are composed as ORNNEIUD ensemble whose prediction accuracy is 63.44%. As shown in Fig.4, ORNNEIUD has no under-fitting individuals and the predictive accuracy of all the individuals are above 56%.

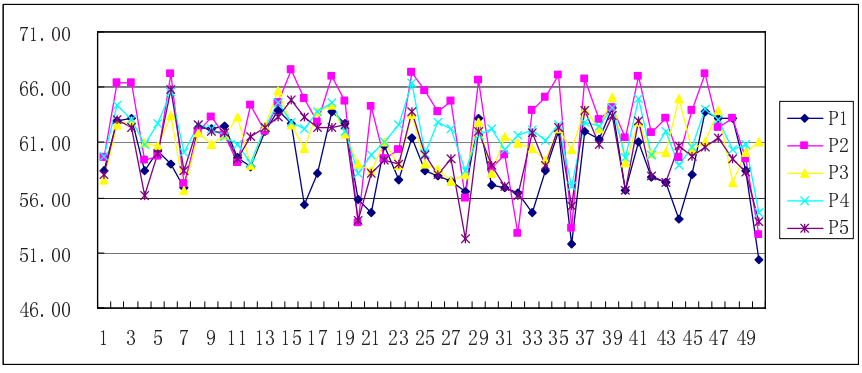


Fig. 3. Predictive accuracy of the “good” individual RBFNNs on the test data set (%)

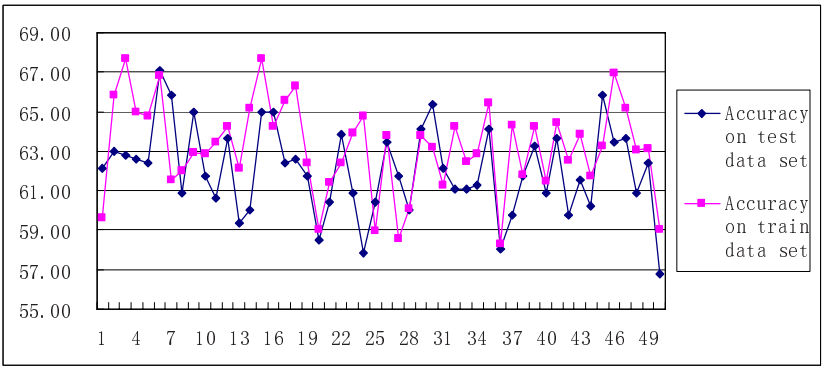


Fig. 4. Predictive accuracy of individual RBFNNs of ORNNEIUD (%)

- **ORNNEUD**. This is also the algorithm we described in section 2 of this paper, in which individual RBFNNs whose prediction accuracies are enhanced after optimization are selected to compose the final ensemble. In the above ORNNEIUD, there are total 17 individual RBFNNs whose performances are enhanced by multi-regression parameter optimization. Now, we select these 17 RBFNNs to constitute the final ensemble, i.e. ORNNEUD.

The final experimental results of the above six different algorithms are listed in Table 3. The comparison of single RBFNN and ORNNUD denotes that even the network parameters are optimized by Uniform Design, single RBFNN is still under-fitting. Bagging method gets a better precision (61.29%) which explains the effectiveness of bagging

method. However, we also observed that 27 out of the 50 individual RBFNNs of bagging ensemble are under-fitting. The generalization accuracy of OBUD is 63.44% and brings a 2 percent improvement of generalization accuracy over the average one of bagging. We also find that bagging is superior to OBUD when the training parameters of individual RBFNN are setup as Parameters 2(63.66%). Compared with bagging, the number of under-fitting RBFNNs also decreases from 27 down to 3. ORNNEIUD doesn't work better than OBUD in prediction accuracy, but has no under-fitting RBFNNs. ORNNEUD performs the best in precision with a 17 percent improvement over single RBFNN, or 5 percent increase over simple bagging ensemble. Meanwhile, the number of individual RBFNNs is reduced to 17 without under-fitting RBFNNs.

Table 3. Experimental results among different algorithms

	Prediction accuracy	Number of individual RBFNNs	Number of under-fitting RBFNNs
RBFNN	49.89%	--	--
ORNNUD	49.89%	--	--
Bagging	61.29%	50	27
OBUD	63.44%	50	3
ORNNEIUD	63.44%	50	0
ORNNEUD	66.24%	17	0

In conclusion, the above experiments results show that:

(1). It's difficult for RBFNN to obtain good performance for QSAR even when uniform design is employed. However, bagging gets better precision which validate the effectiveness of bagging when single neural networks cannot work well.

(2). Bagging by training the individuals with different training parameters gets better results than that with fixed training parameters.

(3). Uniform design improves performance of bagging. However, how to use uniform design is the key problem. Using uniform design to optimize each individual directly just like OBUD is an option. If the individuals trained by the parameter combinations in the uniform design table have poor performance, it is not effective to acquire good individuals by employing re-sampling and uniform design. This is proved by the fact that ORNNEIUD has the same performance with OBUD. Selective ensemble improves the ensemble's performance, which is proved by the results of ORNNEUD.

4 Conclusions

Architecture of an ensemble and training parameters of individual neural networks are critical to generalization performance of an ensemble and convenience of its creation. This paper employs some principles from design of experiment to guide users with little experience of neural networks to designing the ensemble architecture and adjusting the training parameters of individual neural networks and proposed an novel algorithm ORNNEUD (Optimal RBF Neural Network Ensemble based on Uniform Design). At the same time, the nearest-neighbor clustering algorithm is used to create the heterogeneous individuals with different hidden nodes. Furthermore, ORNNEUD

selects the appropriative individual neural networks to constitute the final ensemble according to the experimental results after the uniform design optimization. Both generalization ability of the ensemble and convenience of the ensemble creation are improved through the above mentioned techniques. Experiments on a data set of drug molecular activities demonstrate that accuracy obtained by ORNNEUD is better than that of other ensemble learning methods like bagging and significantly better than that of single RBFNN. ORNNEUD liberates experts from the tedious trial-and-error process to a certain extent when it is applied for QSAR analysis.

In the future, a parallelization strategy for training RBFNN by using MPI (Message Passing Interface) techniques will be proposed to reduce the complexity and improve the computational efficiency of the ensemble's construction. In addition, the idea of ORNNEUD will be applied to make it easily used for our other novel ensemble methods like PREIFEAB [10].

Acknowledgements. This research is supported by the National Natural Science Foundation of China (Grant No. 20503015 and 70502020), National High-Technology Research and Development Program (863 Program) of China (Grant No. 2007AA01Z144), and Shanghai Leading Academic Discipline Project (Grant No. J50103).

References

1. Xu, L., Wu, Y., Hu, C., Li, H.: A QSAR of the Toxicity of Amino-benzenes and Their Structures. Science in China (Series B) 43(2), 130–136 (2000)
2. Chen, N.Y., Lu, W.C., Yang, J., Li, G.Z.: Support Vector Machine in Chemistry. World Scientific Publishing Co. Pt. Ltd., Singapore (2004)
3. Sollich, P., Krogh, A.: Learning with Ensembles: How Over-fitting Can Be Useful. Advances in Neural Information Processing Systems, 190–196 (1996)
4. Slam, M.M., Yao, X., Murase, K.: A Constructive Algorithm for Training Cooperative Neural Network Ensembles. IEEE Trans. on Neural Networks 14(4), 820–834 (2003)
5. Liu, Y., Li, Y., Li, G.Z., Zhang, B.F., Wu, G.F.: Constructive Ensemble of RBF Neural Networks and Its Application to Earthquake Prediction. In: Wang, J., Liao, X.-F., Yi, Z. (eds.) ISNN 2005. LNCS, vol. 3496, pp. 532–537. Springer, Heidelberg (2005)
6. Akhand, M.A.H., Murase, K.: A Minimal Neural Network Ensemble Construction Method: a Constructive Approach. Advanced Computational Intelligence and Intelligent Informatics 11(6), 582–592 (2007)
7. Zhou, Z.H., Wu, J., Tang, W.: Ensembling Neural Networks: Many Could Be Better Than All. Artificial Intelligence 137(1-2), 239–263 (2002)
8. Li, G.Z., Yang, J., Kong, A.S.: Neural Network Ensemble Based on Clustering Algorithm. Journal of Fudan University 5, 689–691 (2004) (in Chinese)
9. Fang, K.T.: Uniform Design and the Uniform Design Table. Science Press, Beijing (1994) (in Chinese)
10. Li, G.Z., Meng, H.H., Lu, W.C., Yang, J.Y., Yang, M.Q.: Asymmetric Bagging and Feature Selection for Activities Prediction of Drug Molecules. BMC Bioinformatics 9 (Suppl. 6), S7 (2008)

Prediction of RNA-Binding Residues in Proteins Using the Interaction Propensities of Amino Acids and Nucleotides

Rojan Shrestha, Jisu Kim, and Kyungsook Han*

School of Computer Science and Engineering, Inha University, Incheon 402-751, Korea
rojan@inhaian,net, sujiper@inhaian.net, khan@inha.ac.kr

Abstract. Recently several machine learning approaches have been attempted to predict RNA-binding residues in amino acid sequences. None of these consider interacting partners (i.e., RNA) for a given protein when predicting RNA-binding amino acids, so they always predict the same RNA-binding residues for a given protein even if the protein may bind to different RNA molecules. In this study, we present a support vector machine (SVM) classifier that takes an RNA sequence as well as a protein sequence as input and predicts potential RNA-binding residues in the protein. The interaction propensity between an amino acid and nucleotide obtained from the extensive analysis of the representative protein-RNA complexes in the Protein Data Bank (PDB) was encoded in the feature vector of the SVM classifier. Four biochemical properties of an amino acid (the side chain pKa value, hydrophobicity index, molecular mass, and accessible surface area) were also encoded in the feature vector. On a dataset of 145 protein sequences and 78 RNA sequences, the SVM classifier achieved a sensitivity of 72.30% and specificity of 78.03%.

Keywords: RNA-binding residues, protein-RNA interactions, support vector machine.

1 Introduction

Protein-RNA interactions are involved in many functions within the cell. For example, tRNAs are bound to aminoacyl-tRNA synthetases for the proper translation of the genetic code into the amino acid of proteins [1] and ribonucleoprotein particles (RNPs) bind RNA in post-transcriptional regulation of gene expression [2]. However, the precise mechanisms of protein-RNA interactions are not fully understood.

A variety of problems concerned with protein-DNA interactions have been investigated [3-8], but protein-RNA interactions have received much less attention despite their importance. This is partly because there were fewer protein-RNA complex structures determined so far than protein-DNA complex structures. However, the number of protein-RNA complex structures has increased substantially in the last few years.

Several machine learning approaches have been applied to predict RNA-binding residues in amino acid sequences. Based on chemical properties of the input protein

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 114 – 121, 2008.
© Springer-Verlag Berlin Heidelberg 2008

sequence, BindN [9] predicts potential RNA or DNA-binding residues with a support vector machine (SVM). RNABindR [10, 11] predicts RNA-binding residues in a protein using a Naïve Bayes classifier. Knowledge of the protein structure is not required. However, none of the previous approaches consider interacting partners (i.e., RNA) of a given protein when predicting RNA-binding amino acids, so they always predict the same RNA-binding residues for a given protein even if the protein may bind to different RNA molecules. As an example of changes in RNA-binding residues due to changes in its binding partners, consider a protein-RNA complex (PDB ID: 1AQ3). Protein chains A and C of the complex have the exactly same amino acid sequence, but they bind to different RNA chains (chains R and S). RNA-binding residues in chains A and C are different from each other.

As an extension of our previous study of protein-RNA interactions [12, 13], we attempt to predict RNA-binding residues in a protein using the interaction propensities of amino acids and nucleotides. In this study, we present an SVM classifier that takes an RNA sequence as well as a protein sequence as input and predicts potential RNA-binding residues in the protein. The interaction propensity between an amino acid and nucleotide, which was obtained from the extensive analysis of the structure data of protein-RNA complexes in the Protein Data Bank (PDB) was encoded in the feature vector of the SVM classifier. Four biochemical properties of an amino acid (the side chain pKa value, hydrophobicity index, molecular mass, and accessible surface area) were also encoded in the feature vector. Experimental results with structure data of protein-RNA complexes show that our SVM classifier achieved a sensitivity of 72.30% and a specificity of 78.03%, which are better than previous studies. Our prediction method enables us to find potential RNA interfaces in proteins and RNA-binding proteins.

2 Materials and Methods

2.1 Dataset of Protein-RNA Complexes

From the Protein Data Bank (PDB) we obtained 253 protein-RNA complexes that had been solved by X-ray crystallography with a resolution of 3.0 Å or better. We executed the blastall program in the BLAST package (http://www.ncbi.nlm.nih.gov/BLAST/download.shtml) on the protein sequences of the complexes with the sequence identity above 90% and the e-value below 0.001 in order to eliminate redundant protein sequences. 75 complexes out of 253 protein-RNA complexes were left, which contain 155 protein sequences and 78 RNA sequences. Protein sequences with no binding RNA sequence were removed further. The final dataset consists of 145 protein sequences and 78 RNA sequences. The dataset has more protein sequences than RNA sequences since there are protein-RNA complexes with multiple protein sequences binding to a single RNA sequence.

2.2 Definition of RNA-Binding Residues

Amino acids in protein-RNA complexes are classified as RNA-binding residues if they are involved in at least one of the three interactions between protein and RNA: hydrogen bonds, van der Waals and hydrophobic interactions. We developed a program for identifying these interactions from PDB files.

Hydrogen bonds (H bonds) are identified by finding all proximal atom pairs between H bond donors (D) and acceptors (A) that satisfy the following geometric criteria: contacts with D-A distance < 3.35 Å, H-A distance < 2.7 Å, and both D-H-A and H-A-AA angles ≥ 90˚, where AA is an acceptor antecedent. van der Waals contacts are the contacts with D-A distance in the range 3.35-3.9 Å, H-A distance ≤7 Å, and both D-H-A and H-A-AA angles ≥ 90˚. Hydrophobic interactions are those with distance between non-polar atoms within 5 Å [14].

2.3 Feature Extraction and Encoding Scheme

We have selected eight different features to encode an amino acid: four biochemical properties of an amino acid (side chain pK_a value, hydrophobicity, molecular mass and accessible surface area) and four weighted interaction propensities for binding with four different types of nucleotides.

The pK_a values of amino acid side chains play an important role in defining the pH-dependent characteristics of a protein. The pH-dependence of the activity displayed by enzymes and the pH-dependence of protein stability, for example, are properties that are determined by the pK_a values of amino acid side chains [16]. Molecular mass of an amino acid gives a unique value so it gives the best approaches to represent a feature vector. Accessible surface area (ASA), hydrophobic scale and side-chain pK_a values play critical roles in molecular interaction because they are either fully or partially participating in ionization process. In this study, we used the ASA value from [18], the hydrophobicity scale developed by Kyte and Doolittle [17] and the side-chain pK_a values from [16].

Since we take both protein and RNA sequences as input, the interaction propensity between amino acids and nucleotides is used for finding RNA-binding amino acids. We define the interaction propensity P_{ab} between amino acid a and nucleotide b by equation (1), where N_{ab} is the number of amino acids a binding to nucleotide b, $\sum N_{ij}$ is the total number of amino acids binding to any nucleotide, N_a is the number of amino acids a, $\sum N_i$ is the total number of amino acids, N_b is the number of nucleotides b, $\sum N_j$ is the total number of nucleotides, and the numbers refer to residues on the surface [12].

$$P_{ab} = \frac{N_{ab}}{\sum N_{ij}} \Big/ \frac{N_a . N_b}{\sum N_i . \sum N_j} \tag{1}$$

Using the formula in equation (1), we calculated the interaction propensity of every combination of amino acids and nucleotides in the protein-RNA complexes in the dataset. The interaction propensity was computed independently for H-bonding interactions, van der Waals contacts and hydrophobic interactions (Fig. 1).

Interestingly, in all three interactions, arginine shows the highest interaction propensity value. On average, arginine (Arg, R) and lysine (Lys, K), the positively charged amino acids, have the highest interaction propensity. This result suggests their consistent ability to participate in interactions both with bases and with the negatively charged phosphate backbone of RNA nucleotides. It also corroborates the interaction propensities obtained from the extensive analysis of 45 protein-RNA complexes in our previous study [12].

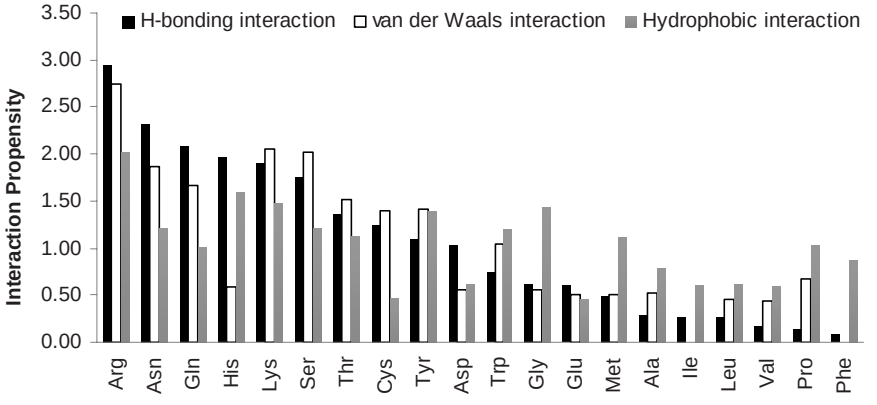


Fig. 1. Interaction propensity of amino acids with RNA in terms of three different types of interactions: H bond interaction, van der Waals contact, and hydrophobic interaction

The features of a target residue and its adjacent residues on left and right sides are represented in a feature vector. When a window of size *w* is used, the number of feature vectors for each data instance is 8*w*. We tested different window sizes, and the window size of 9 (*w*=9) gave the best result. Thus, the total number of feature vectors for each data instance is 72. Each data instance is labeled +1 when the target amino acid is a binding residue and -1 when it is a non-binding residue. Fig. 2 shows an example of a feature vector for an amino acid when a window of size 3 is used.

pK_a (f1)			hydropathy (f2)			mass (f3)			ASA (f4)		
f(1,i-1)	f(1,i)	f(1,i+1)	f(2,i-1)	f(2,i)	f(2,i+1)	f(3,i-1)	f(3,i)	f(3,i+1)	f(4,i-1)	f(4,i)	f(4,i+1)

wIP_A (f5)			wIP_G (f6)			wIP_C (f7)			wIP_U (f8)		
f(5,i-1)	f(5,i)	f(5,i+1)	f(6,i-1)	f(6,i)	f(6,i+1)	f(7,i-1)	f(7,i)	f(7,i+1)	f(8,i-1)	f(8,i)	f(8,i+1)

Fig. 2. Example of a feature vector for amino acid r_i. When a window size of 3 is used, a single amino acid is represented by a 24-element feature vector, 3 elements for each of the 8 properties. wIP_A is the weighted interaction propensity of the amino acid with nucleotide A (adenine). wIP_A is computed by: (interaction propensity of the amino acid with nucleotide A) × (number of A's in the RNA sequence)/(RNA sequence length).

2.4 Bootstrapping for Resampling Negative Data

Out of the total 30,471 amino acids in the dataset, 3,591 amino acids (11.78%) are binding residues, and the remaining 26,880 amino acids (88.22%) are non-binding residues (Fig. 3). There is a severe unbalance between the positive data (set of binding residues) and the negative data (set of non-binding residues). In order to avoid biases towards the overrepresented negative data, we implemented a bootstrapping method [15] to resample the negative data for training. Let $x_i = (f1_i, f2_i, f3_i, f4_i, f5_i, f6_i, f7_i, f8_i)$ be the *i*th feature vector of the negative data instance. A bootstrapping sample

X has *s* feature vectors generated by sampling with replacement *s* times from negative data. The value of *s* was set to 200 in this study. New negative data instances are generated by equation (2). This process is repeated until the number of negative data instances becomes equal to the number of positive data instances.

$$\text{New negative data} = \left(\frac{1}{s}\sum_{i}^{x_i \in X} f1_i, \frac{1}{s}\sum_{i}^{x_i \in X} f2_i, \ldots, \frac{1}{s}\sum_{i}^{x_i \in X} f8_i \right) \tag{2}$$

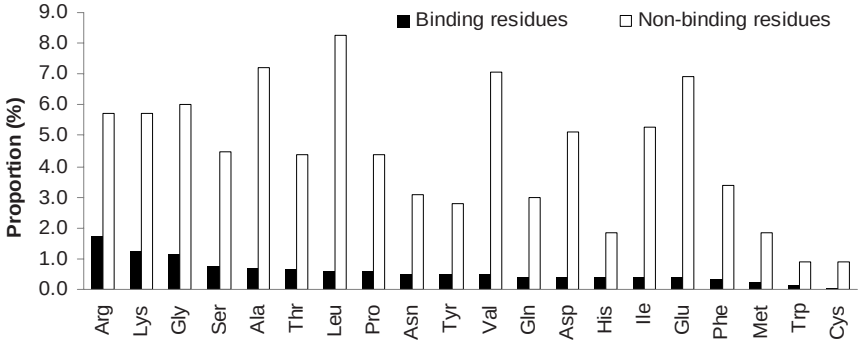


Fig. 3. RNA-binding residues and non-binding residues in the dataset. The horizontal axis represents twenty standard amino acids, and the vertical axis represents the proportion of each amino acid in total amino acids in the dataset.

3 Results and Discussion

In our study, a library for support vector machines (LIBSVM) available at http://www.csie.ntu.edu.tw/~cjlin/libsvm/ was used to construct the SVM classifier model as well as to test the model. As a kernel for the SVM classifier, we used the radial basis function (RBF), which is defined by

$$f(u, v) = \exp(-\gamma \|u - v\|^2) \tag{3}$$

where *u* and *v* are two data vectors and γ is the training parameter.

We randomly partitioned whole data instances into training and testing sets with the ratio 7:3. We constructed several SVM classifiers by varying parameter values. The value of γ in equation (3) was computed by taking the inverse of the size of a feature vector. For a window size of 3, an amino acid is represented by a 24-element feature vector, and γ =1/24.

After running the SVM model on the test data set, we counted true positives, true negatives, false positives, and false negatives by comparing the prediction results with the actual data.

- True positives (TP): binding residues that are predicted as binding residues
- True negatives (TN): non-binding residues that are predicted as non-binding residues
- False positives (FP): non-binding residues that are predicted as binding residues
- False negatives (FN): binding residues that are predicted as non-binding residues

We used sensitivity and specificity in addition to accuracy to evaluate the SVM model. Sensitivity is the percentage of residues that are RNA-binding and are correctly predicted as RNA-binding. Specificity is the percentage of residues that are not RNA-binding and are correctly predicted as non-binding. Accuracy is the percentage of residues that are correctly predicted.

$$Sensitivity = \frac{TP}{TP + FN} \tag{4}$$

$$Specificity = \frac{TN}{TN + FP} \tag{5}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

Table 1 shows the performance of the SVM classifier with different window sizes. Window sizes do not make a big difference in the performance, but the SVM classifier with a window size of 9 gave the best result in terms of accuracy.

Table 1. Performance of the SVM classifier with different window sizes for the dataset of 145 protein sequences and 78 RNA sequences in section 2.1

window size	accuracy (%)	sensitivity (%)	specificity (%)
5	73.44	67.04	76.56
7	73.77	68.13	76.53
9	75.92	72.30	78.03
11	75.91	72.50	77.66
13	75.63	72.62	77.20
15	75.89	70.82	77.37

As an example of predicting binding sites, Fig. 4 shows a tertiary structure of a protein-RNA complex (PDB ID: 20IH). For this test case, the SVM classifier achieved a sensitivity of 81.48% and specificity of 79.41%.

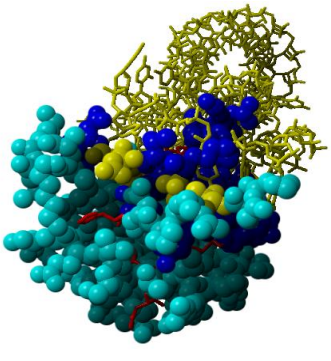


Fig. 4. Prediction results for a protein-RNA complex (PDB ID: 20IH). 22 binding residues (shown as blue balls) and 54 non-binding residues (cyan balls) are predicted correctly. 14 residues (red wireframe) are not actually binding but predicted as binding. 5 residues (yellow balls) are acutally binding but predicted as non-binding. RNA is shown as a yellow wireframe.

To further evaluate our SVM classifier, we tested it on two additional datasets: (1) what we call the PRNA-47 dataset of 47 non-redundant protein sequences and 47 RNA sequences, which was collected in our study, and (2) the PRINR25 dataset of BindN [9]. Our SVM classifier achieved 63.28% sensitivity and 97.88% specificity on the PRNA-47 dataset. On the PRINR25 dataset, it achieved 31.17% sensitivity and 94.75% specificity. BindN showed 66.28% sensitivity and 69.84% specificity on the PRINR25 dataset [9], so our classifier showed a lower sensitivity but a higher specificity than BindN. However, it should be noted that the performance of BindN was achieved by 5-fold cross validation on the PRINR25 dataset, whereas the dataset was not used for training our SVM classifier but only for testing it.

4 Conclusion

In this paper, we presented a SVM classifier that takes both RNA and protein sequences as input and predicts potential RNA-binding residues in the protein. The interaction propensity between an amino acid and nucleotide, which was obtained from the extensive analysis of the structure data of protein-RNA complexes in the Protein Data Bank (PDB) was encoded in the feature vector of the SVM classifier. Four biochemical properties of an amino acid (the side chain pKa value, hydrophobicity index, molecular mass, and accessible surface area) were also encoded in the feature vector. Experimental results with representative protein-RNA complexes show that our SVM classifier achieved a sensitivity of 72.30% and a specificity of 78.03% on the dataset of 145 protein sequences and 78 RNA sequences.

As a future work, we plan to predict protein-binding nucleotides in RNA as well. The problem of predicting protein-binding nucleotides in RNA is much harder to solve than its inverse problem (i.e., predicting RNA-binding residues in protein). One reason for this is because nucleotides are much less diverse than amino acids in their interaction propensities. A method that is capable of predicting both protein-binding nucleotides in RNA and RNA-binding residues in protein will provide useful information for designing new drugs for RNA-binding protein targets or protein-binding RNA targets.

Acknowledgments. This work was supported by the KOSEF grant (No. R01-2008-000-10197-0) funded by the Korea government (MEST).

References

1. Moras, D.: Aminoacyl-t RNA Synthetases. Curr. Opin. Struct. Biol. 2, 138–142 (1992)
2. Varani, G., Nagai, K.: RNA Recognition by RNP Proteins During RNA Processing. Annu. Rev. Biophys. Biomol. Struct. 27, 407–445 (1998)
3. Jones, S., van Heyningen, P., Berman, H.M.: Protein-DNA Interactions: A Structural Analysis. Mol. Biol. 287, 877–896 (1999)
4. Pabo, C.O., Nekludova, L.: Geometric Analysis and Comparison of Protein-DNA Interfaces: Why is There No Simple Code for Recognition? J. Mol. Biol. 301, 597–624 (2000)

5. Ofran, Y., Mysore, V., Rost, B.: Prediction of DNA-binding Residues from Sequence. Bioinformatics 23, 347–353 (2007)
6. Tjong, H., Zhou, H.X.: DISPLAR: An Accurate Method for Predicting DNA-binding Sites on Protein Surfaces. Nucleic Acids Research 35, 1465–1477 (2007)
7. Hwang, S., Gou, Z., Kuznetsov, I.B.: DP-Bind: A Web Server for Sequence-based Prediction of DNA-binding Residues In DNA-binding Proteins. Bioinformatics 23, 634–636 (2007)
8. Ahmad, S., Gromiha, M.M., Sarai, A.: Analysis and Prediction of DNA-binding Proteins and Their Binding Residues Based on Composition, Sequence and Structural Information. Bioinformatics 20, 477–486 (2004)
9. Wang, L., Brown, S.J.: BindN: A Web-based Tool for Efficient Prediction of DNA and RNA Binding Sites in Amino Acid Sequences. Nucleic Acids Research 34 Web Server issue, 243–248 (2006)
10. Terribilini, M., Sander, J.D., Lee, J.H., Zaback, P., Jernigan, R.L., Honavar, V., Dobbs, D.: RNABindR: A Server for Analyzing and Predicting RNA-binding Sites in Proteins. Nucleic Acids Research 294, 1–7 (2007)
11. Terribilini, M., Lee, J., Yan, C., Jernigan, R.L., Honavar, V., Dobbs, D.: Prediction of RNA Binding Sites in Proteins from Amino Acid Sequence. RNA 12, 1450–1462 (2006)
12. Han, K., Nepal, C.: PRI-Modeler: Extracting RNA Structural Elements from PDB Files of Protein-RNA Complexes. FEBS Letters 581, 1881–1890 (2007)
13. Kim, H., Jeong, E., Lee, S.-W., Han, K.: Computational Analysis of Hydrogen Bonds in Protein-RNA Complexes for Interaction Patterns. FEBS Letters 552, 231–239 (2003)
14. Allers, J., Shamoo, Y.: Structure-based Analysis of Protein-RNA Interactions using the Program ENTANGLE. J. Mol. Biol. 311, 75–86 (2001)
15. Dupret, G., Koda, M.: Bootstrap Re-sampling for Unbalanced Data in Supervised Learning. European Journal of Operational Research 134, 141–156 (2001)
16. Nelson, D.L., Cox, M.M.: Lehninger Principles of Biochemistry, p. 3. Worth Publishers, New York (2000)
17. Kyte, J., Doolittle, R.F.: A Simple Method for Displaying the Hydropathical Character of a Protein. J. Mol. Biol. 157, 105–132 (1982)
18. Chothia, C.: The Nature of the Accessible and Buried Surface in Proteins. J. Mol. Biol. 105, 1–14 (1976)

EMD Approach to Multichannel EEG Data - The Amplitude and Phase Synchrony Analysis Technique

Tomasz M. Rutkowski[1], Danilo P. Mandic[2], Andrzej Cichocki[1], and Andrzej W. Przybyszewski[3,4]

[1] Laboratory for Advanced Brain Signal Processing
RIKEN Brain Science Institute, Japan
tomek@brain.riken.jp
http://www.bsp.brain.riken.jp/
[2] Imperial College London, United Kingdom
d.mandic@imperial.ac.uk
http://www.commsp.ee.ic.ac.uk/~mandic/
[3] Department of Psychology, McGill University, Montreal, Canada
przy@ego.psych.mcgill.ca
[4] Department of Neurology, University of Massachusetts Medical School, MA, USA
http://www.umassmed.edu/neurology/faculty/Przybyszewski.cfm

Abstract. Human brains expose the possibility to be connected directly to the intelligent computing applications in form of brain computer/machine interfacing (BCI/BMI) technologies. Neurophysiological signals and especially electroencephalogram (EEG) are the forms of brain electrical activity which can be easily captured and utilized for BCI/BMI applications. Those signals are unfortunately highly contaminated by noise due to a very low level of electrophysiological signals and presence of different devices in the environment creating electromagnetic interference. In the proposed approach we first decompose each of the recorded channels, in multichannel EEG recording environment, into intrinsic mode functions (IMF) which are a result of empirical mode decomposition (EMD) extended to multichannel analysis in this paper. We present novel and interesting results on human mental and cognitive states estimation based on analysis of the above mentioned stimuli-related IMF components.

Keywords: EEG; brain synchrony; brain signal processing; EMD application to EEG.

1 Introduction

Online brain states analysis based on non-invasive monitoring techniques such us EEG have received much attention recently due to the growing interest and popularity of research related to BCI/BMI techniques, owing to the very exciting possibility of computer-aided communication with the outside world. Also a new and growing interest in neuroscience in so called steady-state potentials [1,2,3], which produce longer in time and more easy to detect within monitored EEG

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 122–129, 2008.
© Springer-Verlag Berlin Heidelberg 2008

steady responses contribute also to EEG signal processing popularity. EEG based brain stages monitoring is achieved in a non-invasive recording setup, which poses several important and difficult problems. In terms of signal processing these include the detection, estimation, interpretation and modeling of brain activities, and cross-user transparency [4]. It comes as no surprise, therefore, that this technology is envisaged to be at the core of future "intelligent computing". Other industries which would benefit greatly from the development of online analysis and visualization of brain states include the prosthetics, entertainment, and computer games industries, where the control and navigation in a computer-aided application is achieved without resorting to using hands, or gestures (peripheral nervous system in general). Instead, the onset of "planning an action" recorded from the head (scalp) surface, and the relevant information is "decoded" from this information carrier.

Apart from purely signal conditioning problems, in most BCI/BMI experiments other issues such us user training and adaptation, which inevitably causes difficulties and limits in wide spread of this technology due to the lack of "generality" caused by cross-user differences [4]. To help mitigate some of the above mentioned issues, we propose to make use of a new and growing in interest in signal processing community technique of Empirical Mode Decomposition (EMD) [5] which we extend to multichannel approach of parallel decomposition of single channel signals and further comparison of obtained components among channels to track coherent (synchronized) activities in complex signals as EEG.

In the proposed approach, we analyze responses from experiments based on a visual stimuli which were conducted with in Laboratory for Advanced Brain Signal Processing, BSI RIKEN, within the so-called Steady State Visual Evoked Potential (SSVEP) mode [1,6]. Within this framework, the subjects are asked to focus their attention on simple flashing stimuli, whose frequency is known to cause a physiologically stable response traceable within EEG [6,7]. This way, the proposed multichannel and multimodal signal decomposition scheme uses the EEG captured by several electrodes, subsequently preprocessed, and transformed into informative time-frequency traces, which very accurately visualize frequency and amplitude modulations of the original signal.

EEG is usually characterized as a summation of post-synaptic potentials from a very large number of neurons which create oscillatory patterns distributed and possible to record around the scalp. Those patterns in the known frequency ranges can be monitored and classified in synchrony to a stimuli given to the subjects. EMD utilizes empirical knowledge of oscillations intrinsic to a time series in order to represent them as a into superposition of components with well defined instantaneous frequencies. These components are called intrinsic mode functions (IMF).

This paper is organized as follows. First the method of single channel EMD analysis of EEG signals is presented. Next multichannel EEG analysis and decomposition is discussed leading to a novel time–frequency synchrony evaluation method. A new concept of multiple spatially localized amplitude and frequency oscillations related to presented stimuli in time-frequency domain is descirbed

which let us obtain final traces of frequency and amplitude ridges. Finally examples of the analysis of the EEG signals are given and conclusions are drawn.

2 Methods

We aim at looking at the level of detail (richness of information source) obtainable from single experimental trial EEG signals, and compare the usefulness of a novel multichannel signal decomposition approach in this context. The utilized approach is based on multichannel extension of the EMD technique which was previously successfully applied to EEG sonification in [1,2,3]. The EMD approach rests on the identification of signal's non-stationary and non-linear features which represent different modalities of brain activity captured by the EEG data acquisition system (g.USBamp® of Guger Technologies). This novel method allowed us previosly to create slowly modulated tones representing changing brain activities among different human scalp locations where EEG electrodes were localized.

In the current application we propose to look at the level of details revealed in single EEG channels decomposed separately into IMFs as discussed in detail in the next section and later compared across the multiple channels. After application of Hilbert transform to all IMFs we can track and visualize revealed oscillations of amplitude and frequency ridges. Similarity of these oscillations among channels revealed in form of pairwise correlations identify components which are synchronized or not with the onsets and offsets of the presented stimuli. The proposed approach is the completely "data driven" concept. All the IMFs create semi-orthogonal bases created from the original EEG signals and not introduced artificially by the method itself.

2.1 Empirical Mode Decomposition for EEG Analysis - A Single Channel Case

The IMF components obtained during EMD analysis should approximately obey the requirements of (i) completeness; (ii) orthogonality; (iii) locality; (iv) adaptiveness [5]. To obtain an IMF it is necessary to remove local riding waves and asymmetries, which are estimated from local envelope of minima and maxima of the waveform. There are several approaches to estimate signals envelopes and we discussed them previously [3].

The Hilbert spectrum for a particular IMF allows us later to represent the EEG in amplitude - instantaneous frequency - time plane. An IMF shall satisfy the two conditions: (i) in the whole data set, the number of extrema and the number of zero crossings should be equal or differ at most by one; (ii) at any point of IMF the mean value of the envelope defined by the local maxima and the envelope defined by the local minima should be zero.

The technique of finding IMFs corresponds thus to finding limited-band signals. It also corresponds to eliminating riding-waves from the signal, which ensures that the instantaneous frequency will not have fluctuations caused by an asymmetric wave form. IMF in each cycle is defined by the zero crossings. Every IMF involves

only one mode of oscillation, no complex riding waves are thus allowed. Notice that the IMF is not limited to be a narrow band signal, as it would be in traditional Fourier or wavelets decomposition, in fact, it can be both amplitude and frequency modulated at once, and also non-stationary or non-linear.

The process of IMF extraction from a signal $x(t)$ ("sifting process" [5]) is based on the following steps:

1. determine the local maxima and minima of the analyzed signal $x(t)$;
2. generate the upper and lower signal envelopes by connecting those local maxima and minima respectively by the chosen interpolation method (e.g., linear, spline, cubic spline, piece-wise spline [3,5]);
3. determine the local mean $m(t)$, by averaging the upper and lower signal envelopes;
4. subtract the local mean from the data: $h_1(t) = x(t) - m_1(t)$.

Ideally, $h_1(t)$ is an IMF candidate. However, in practice, $h_1(t)$ may still contain local asymmetric fluctuations, e.g., undershoots and overshoots; therefore, one needs to repeat the above four steps several times, resulting eventually in the first (optimized) IMF. In order to obtain the second IMF, one applies the sifting process to the residue $\varepsilon_1(t) = x(t) - \mathrm{IMF}_1(t)$, obtained by subtracting the first IMF from $x(t)$; the third IMF is in turn extracted from the residue $\varepsilon_2(t)$ and so on. One stops extracting IMFs when two consecutive sifting results are close to identical; the empirical mode decomposition of the signal $x(t)$ may be written as:

$$x(t) = \sum_{k=1}^{n} \mathrm{IMF}_k(t) + \varepsilon_n(t), \tag{1}$$

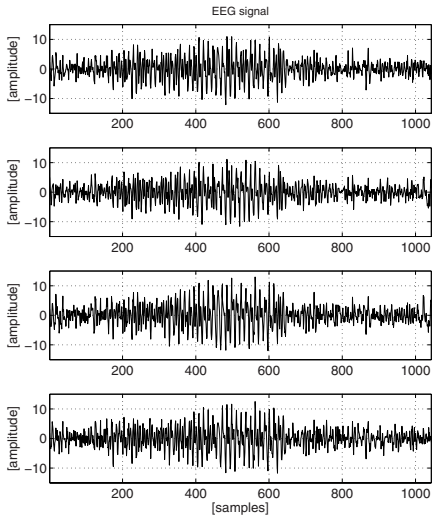
where n is the number of extracted IMFs, and the final residue $\varepsilon_n(t)$ can either be the mean trend or a constant. The EMD is obviously complete, since (1) is an equality: the original signal can be reconstructed by adding all IMFs and the final residue. Note that the IMFs are not guaranteed to be mutually orthogonal, but in practice, they often are close to orthogonal [3]; it is also noteworthy that the IMFs are adaptive, i.e., they depend on the signal $x(t)$ as anticipated for the data driven method.

2.2 Hunang-Hilbert Spectra with Amplitude and Frequency Ridges

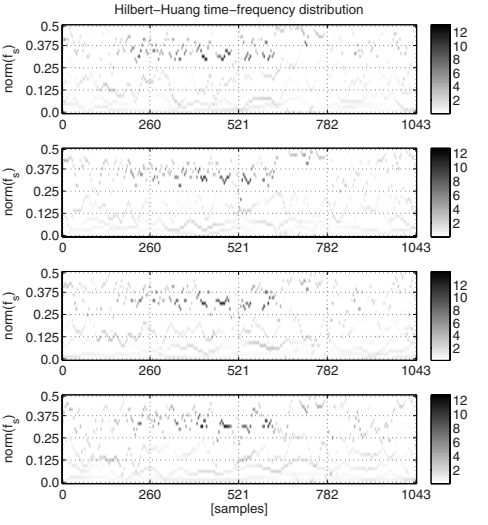
From the obtained in previous section IMFs corresponding time–frequency representations can be produced by applying the Hilbert transform to each component [5]. As a result of Hilbert transform application to each IMF the data can be expressed time-frequency domain as:

$$R(t) = \sum_{k=1}^{n} \mathrm{IMF}_k(t) \exp\left(\mathrm{i} \int \omega_k(t) dt\right), \tag{2}$$

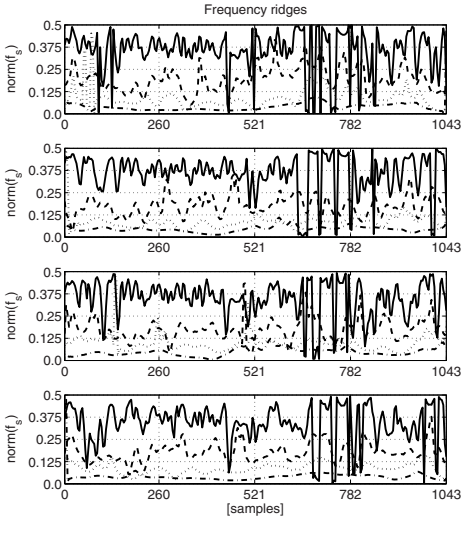
where IMF_k are obtained as above discussed and ω_j is an instantaneous frequency. The Hilbert transform allows us to depict the variable amplitude (Figure 1(d)) and



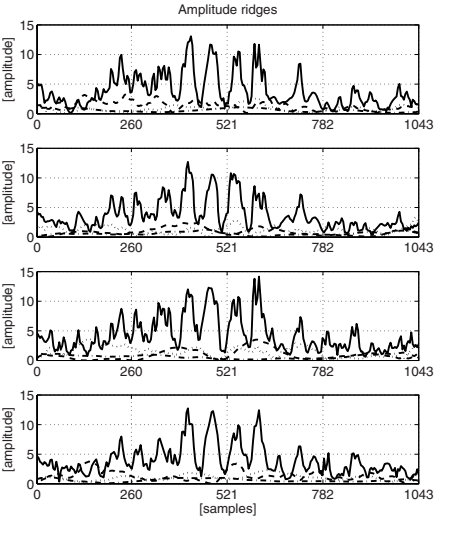
(a) Time domain preprocessed EEG.



(b) Huang-Hilbert spectra.



(c) Frequency ridge traces.



(d) Amplitude ridge traces

Fig. 1. Four plots of four EEG channels in each panel recorded synchronously during SSVEP experiment. The steady-state response can be visually spotted in the range of $150-650$ samples. Panel (a) presents time domain EEG preprocessed plots; (b) their Huang-Hilbert spectra; (c) and (d) the frequency and amplitude ridges in Hilbert spectra domain (solid line: first; dashed line: second; dotted line:third; dash-dotted line: fourth IMF respectively).

the instantaneous frequency (Figure 1(c)) in the form of very sharp and localized functions of frequency and time (in contrast to Fourier expansion, for example, where frequencies and amplitudes are fixed for its bases). Such an approach is very suitable for the non-stationary EEG analysis and common/sychronized activities within certain channels. An example of Huang-Hilbert spectrograms of four EEG channels recorded simultaneously is presented in Figure 1(b).

2.3 EMD Application to to Multichannel EEG Signals

Using the above procedure in a single channel mode the EEG signals from chosen electrodes could be decomposed separately forming subsets of IMF functions, from which low frequency drifts and high frequency spikes could be further removed. The most interesting part of EEG is usually in the middle range frequencies. To analyze multichannel EEG signal sets recorded synchronously in a single experiment we propose to decompose all channels separately preventing possible oscillatory information leaking among the channels. The so obtained IMFs sets can be further

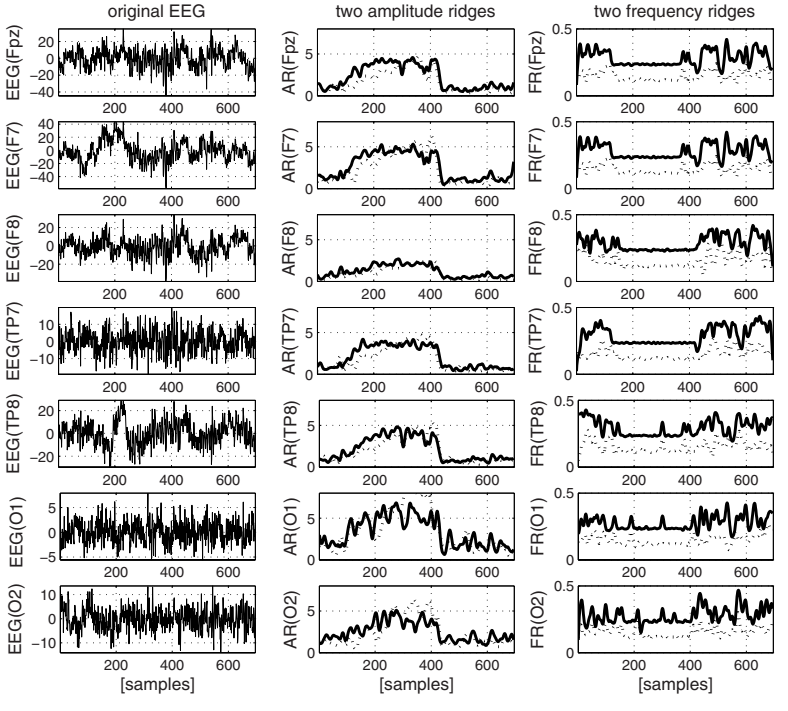


Fig. 2. The result showing the power of the proposed method to analyze multichannel EEG recordings. The first column shows noisy EEG signals, while the second and third columns depict only amplitude (AR) and frequency (FR) traces of components synchronized with the stimuli and subsequently correlated among channels (solid line for the first and dashed line for the second IMF synchronized with the stimuli).

compared as in case of four EEG signals presented in Figure 1(a) which are further EMD decomposed and visualized in form of Huang-Hilbert spectrograms [5] as in Figure 1(b). The traces of amplitude and frequency modulations ridges [8] obtained from Hilbert transformation of separately processed IMFs and further plotted together are presented in Figures 1(d) and 1(c) respectively. Ridges are the continues traces within spectrograms of frequency and amplitude oscillations as first introduced in [8]. The areas of steady-state stimulation can be easily spotted in amplitude traces of a single IMF on all channels in Figure 1(d) and subsequently in form of stable frequency ridge during stimulation with very strong oscillations before and after the stimuli in all channels as in Figure 1(c).

The combined result of analysis of seven EEG channels from locations around the human head during similar SSVEP stimuli BCI/BMI paradigm is shown also in Figure 2. There are seven time domain EEG traces plotted with only two amplitude (AR) and frequency (FR) ridges of the only components showing synchrony with the stimuli. Those components were chosen based on cross-channel correlation of IMFs showing again that the proposed multichannel EMD analysis is useful for spotting steady-state responses from noisy EEG based on amplitude and frequency modulations synchrony.

3 Conclusions

A new approach to extend a single channel EMD approach to EEG signals analysis with steady-state responses perfectly suited for further BCI/BMI application is presented. We propose to decompose EEG signals separately to obtain sets of IMF components. The proposed analysis of multichannel Hilbert-Huang domain frequency and amplitude ridges together with their subsequent correlation among channels during the steady-state stimuli presentation revealed a possibility to identify different brain states related to the stimuli. This allows us to detect from a very noisy and single trial EEG signals the time slots when stimuli was presented to the subjects. Such approach is perfectly suited to steady-state stimuli based BCI/BMI applications where online markers of users changes of attention to different external stimuli is sought.

Acknowledgements. This work was supported in part by JSPS and Royal Society under the Japan-UK Research Cooperative Program, 2007 & 2008.

References

1. Rutkowski, T.M., Vialatte, F., Cichocki, A., Mandic, D., Barros, A.K.: Auditory Feedback for Brain Computer Interface Management - An EEG Data Sonification Approach. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS (LNAI), vol. 4253, pp. 1232–1239. Springer, Heidelberg (2006)
2. Rutkowski, T.M., Cichocki, A., Ralescu, A.L., Mandic, D.P.: Emotional States Estimation from Multichannel EEG Maps. In: Wang, R., Gu, F., Shen, E. (eds.) Advances in Cognitive Neurodynamics ICCN 2007 - Proceedings of the International Conference on Cognitive Neurodynamics. Neuroscience, Springer, Heidelberg (in press, 2008)

3. Rutkowski, T.M., Cichocki, A., Mandic, D.P.: EMDsonic - An Empirical Mode Decomposition Based Brain Signal Activity Sonification Approach. Information Technology: Transmission, Processing and Storage. In: Signal Processing Techniques for Knowledge Extraction and Information Fusion, February 2008, pp. 261–274. Springer, Heidelberg (2008)
4. Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M.: Brain-computer Interfaces for Communication and Control. Clinical Neurophysiology 113, 767–791 (2002)
5. Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Yen, N.C., Tung, C., Liu, H.: The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis. In: Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, March 1998, vol. 454, pp. 903–995 (1998)
6. Kelly, S.P., Lalor, E.C., Finucane, C., McDarby, G., Reilly, R.B.: Visual Spatial Attention Control in an Independent Brain-computer Interface. IEEE Transactions on Biomedical Engineering 52(9), 1588–1596 (2005)
7. Niedermeyer, E., Da Silva, F.L. (eds.): Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 5th edn. Lippincott Williams & Wilkins (2004)
8. Przybyszewski, A.W., Rutkowski, T.M.: Processing of the Incomplete Representation of The Visual World. In: Hryniewicz, O., Kacpryzk, J., Koronacki, J., Wierzchon, S. (eds.) Issues in Intelligent Systems Paradigms. Problemy Wspolczesnej Nauki - Teoria i Zastosowania - Informatyka, pp. 225–235. Akademicka Oficyna Wydawnicza EXIT, Warsaw, Poland (2005)

Prediction of Binding Sites in HCV Protein Complexes Using a Support Vector Machine

Taihui Yoo, Jaetak Lee, and Kyungsook Han*

School of Computer Science and Engineering, Inha University, Inchon 402-751, Korea
thyoo@inhaian.net, jaetaklee@inhaian.net, khan@inha.ac.kr

Abstract. Hepatitis C virus (HCV) infection is a major cause of liver disease and a dangerous threat to public health. Hence, the problem of interactions between HCV and human proteins has received much attention. In the present study, we propose a support vector machine (SVM) model for predicting the binding residues in HCV protein complexes. The SVM model achieved an average sensitivity of 76.06% and specificity of 75.94% for 18 non-redundant HCV protein complexes. This approach can efficiently search potential protein-binding sites in proteins and a wide range of protein-protein interaction sites.

Keywords: hepatitis C virus, binding sites, support vector machine.

1 Introduction

Hepatitis C virus (HCV) is a small enveloped virus with a single-stranded RNA genome encoding a single open reading frame. The polyprotein of approximately 3,100 amino acids (aa) is cleaved into structural polypeptides (core, E1 and E2) and non-structural polypeptides (NS1, NS2, NS3, NS4A, NS4B, NS5A and NS5B) [1, 2].

Although many experimental studies have been performed so far, the underlying mechanisms controlling the entry of HCV into host cells and interactions with the host cells are not fully known, and an efficient treatment for HCV infection has not been developed. Experimental studies of the interactions between HCV and human proteins are not easy partly because HCV does not grow readily in tissue culture, and until recently, there has been no small animal model to propagate the HCV proteins [3, 4]. Computational approaches have also been attempted to study the interactions between HCV proteins and human proteins. However, it is not easy to identify the patterns of interactions between HCV and human proteins due to a lack of structure data. Thus, identifying the binding sites of HCV protein complexes is a meaningful intermediate step toward understanding the HCV infection mechanism and interactions between HCV and human proteins.

In our previous study [7], we developed a radial basis function neural network (RBFNN) model for predicting the binding sites of HCV protein complexes. Each residue is represented by a 39-bit vector based on its residue type, sequence entropy, binding propensity, biochemical property, and secondary structure information. The

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 130 – 137, 2008.
© Springer-Verlag Berlin Heidelberg 2008

model predicts binding sites of HCV protein complexes with a reasonably good sensitivity of 77% but with a low specificity of 3%.

Motivated by the need for a better method, we have recently developed a support vector machine (SVM) classifier for predicting binding sites of HCV protein complexes. It uses four biochemical properties of an amino acid (hydrophobicity, accessible surface area, residue interface propensity, and protrusion) to identify surface patches with potential binding residues, and encodes the sequence profile and entropy in a feature vector. The SVM classifier achieved an overall sensitivity of 76.06% and specificity of 75.94% for predicting binding residues in non-redundant HCV protein complexes extracted from the Protein Data Bank (PDB) [5]. The rest of the paper describes the method and its prediction results in more detail.

2 Materials and Methods

2.1 Data Set of HCV Protein Complexes

We first retrieved a total of 64 HCV protein complexes from PDB, and executed the PSI-BLAST program [6] to remove proteins with the sequence identity of 90% or higher. 18 non-redundant HCV protein complexes were left in the data set (Table 1). The data set contains a total of 8,821 residues (783 binding residues and 8,038 non-binding residues), and was used to train and test the SVM classifier.

Table 1. Data set of HCV protein complexes

PDB code	Chain	Chain length
1CSJ	A/B	531/531
1NHU	A/B	578/578
1A1R	A/B, C/D	198/198, 23/23
1C2P	A/B	576/576
1HEI	A/B	451/451
1N64	H/L, P	220/218, 16
1NHV	A/B	578/578
1ZH1	A/B	163/163
2AWZ	A/B	580/580
2AX1	A/B	580/580
2D3U	A/B	570/570
2D41	A/B	570/570
2GC8	A/B	578/578
2HWH	A/B	578/576
2I1R	A/B	576/576
1RTL	A/B, C/D	200/200, 23/23
1NS3	A/B, C/D	186/186, 14/14
2A4R	A/B, C/D	200/23, 200/23

2.2 Surface Patch and Feature Vector

In general, protein binding sites are more hydrophobic, planar, globular, and protruding than other parts of the protein surface [11, 12]. In this study, we used a program

called SHARP^2 (http://www.bioinformatics.sussex.ac.uk/SHARP2/ sharp2.html) [12] to generate the parameters for the SVM classifier. SHARP^2 predicts potential protein-protein interaction sites on the surface of proteins by implementing a simple predictive algorithm that calculates multiple parameters for overlapping patches of residues on the surface of a protein [13]. SHARP^2 calculates up to six parameters (solvation potential, hydrophobicity, accessible surface area, residue interface propensity, planarity and protrusion) for each patch, and predicts the patch with the highest combined score as a potential protein-protein interaction site. Out of the six parameters, we use four parameters (hydrophobicity, accessible surface area, residue interface propensity and protrusion) since solvation potential and planarity are redundant with other parameters. We executed SHARP^2 with the four parameters on each of the HCV protein complexes, and identified surface patches with binding sites. Fig. 1 shows an example of a surface patch of 16 residues present in a protein complex.

Residue number	Residue
85	Val
86	Glu
101	Phe
102	Gly
111	Leu
112	Ser
113	Ser
114	Lys
116	Val
117	Asn
118	His
120	His
121	Ser
122	Val
123	Trp
124	Lys

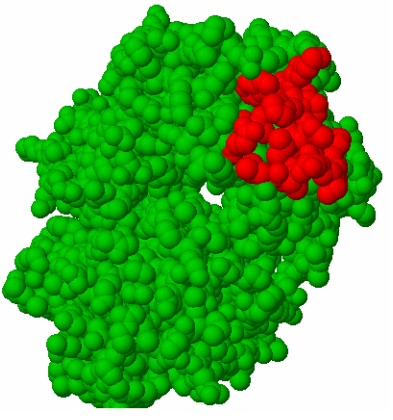


Fig. 1. A surface patch constructed by SHARP^2 for 16 amino acids (red balls) present in HCV protein complex (PDB ID: 1CSJ)

From the surface patches for HCV protein complexes generated by SHARP^2, we selected those also present in the HSSP database [9]. The HSSP database is a derived database with structure and sequence data. For each protein of known 3D structure from PDB, HSSP provides a multiple sequence alignment of putative homologues and a sequence profile of the protein family, centered on the known structure [10].

For the surface patches of the 18 HCV protein complexes, we generated feature vectors using the sequence profile and absolute entropy provided by HSSP. Each amino acid is represented by a 21-element feature vector: 20 elements for the sequence profile and 1 element for the entropy. The sequence profile shows the relative frequencies of 20 amino acids at the position, and the entropy is a measure of sequence variability at the position [8, 9].

2.3 Re-Sampling Data and Classification by SVM

The data set of 18 HCV protein complexes contains 8,038 non-binding residues (91.1% of the total residues) and 783 binding residues (8.9%). Since the negative data set of non-binding residues is much larger than the positive data set of binding residues, prediction by a classifier trained with unbalanced data sets can be biased towards the overrepresented negative data. We implemented a bootstrap method [14, 15] to resample non-binding residues, and constructed a negative data set of comparable size to a positive data set.

The bootstrap method was originally introduced to estimate the standard error of empirical distribution. It allows estimating significant levels of arbitrary statistics when the form of the underlying distribution is unknown [15]. The plug-in estimate of the mean is as follows.

$$\hat{\theta} = E_{\hat{F}}(O^d) = \frac{1}{N}\sum_{i=1}^{N}(O_i^d)^* \quad (1)$$

We used a program called winSVM available at http://www.cs.ucl.ac.uk/staff/M.Sewell/winsvm/ to construct the SVM classifier. In the SVM classifier we tried four kernel functions: polynomial function, radial basis function (RBF), ANOVA function and neural function. Fig. 2 shows the procedure for predicting binding residues using the SVM classifier.

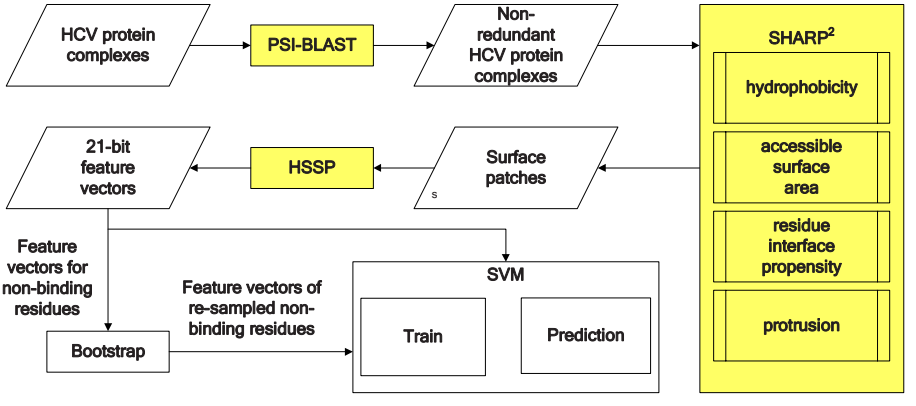


Fig. 2. Procedure for predicting binding sites in HCV protein complexes. A shaded box represents a program or database developed by others.

3 Prediction Results by the SVM Classifier

We tested the SVM classifier on the HCV protein complexes by one-leave-out validation, and computed the sensitivity and specificity of the SVM classifier by equations (5) and (6). Sensitivity is the percentage of residues that are binding and are correctly predicted as binding. Specificity is the percentage of residues that are not binding and are correctly predicted as non-binding.

$$sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$specificity = \frac{TN}{TN + FP} \tag{6}$$

where

- True positives (TP): binding residues that are predicted as binding residues
- True negatives (TN): non-binding residues that are predicted as non-binding residues
- False positives (FP): non-binding residues that are predicted as binding residues
- False negatives (FN): binding residues that are predicted as non-binding residues

Table 2. Prediction results for the 18 HCV protein complexes. The denominators in the second and third columns are the total numbers of binding and non-binding residues in the complex.

PDB code	Binding residues predicted correctly	Non-binding residues predicted correctly	Sensitivity (%)	Specificity (%)
1CSJ	26/33	377/498	78	75
1NHU	24/33	389/525	72	74
1A1R	32/44	98/124	72	79
1C2P	26/33	401/525	78	76
1HEI	50/66	612/789	75	77
1N64	63/79	280/374	79	74
1NHV	28/33	368/525	84	70
1ZH1	48/66	215/260	72	82
2AWZ	25/33	400/524	75	76
2AX1	25/33	410/521	75	78
2D3U	29/33	427/525	87	81
2D41	24/33	398/524	72	75
2GC8	26/33	372/524	78	70
2HWH	25/33	391/526	75	74
2I1R	24/33	403/527	72	76
1RTL	52/66	261/304	78	85
1NS3	25/33	107/139	75	76
2A4R	48/66	210/304	72	69
Average			76.06	75.94

As shown in Table 2, the SVM classifier showed an average sensitivity of 76.06% and specificity of 75.94%. Fig. 3 shows a Receiver Operating Characteristic (ROC) curve for the average prediction results. The ROC curve shows the tradeoff between sensitivity and specificity. In general, the closer the ROC curve to the top left corner, the more accurate the prediction method [16]. The ROC curve for the prediction exhibits a relatively good performance.

Fig. 3. The ROC curve, which shows true positive rate (sensitivity) against the false positive rate (1-specificity)

As an example of predicting binding sites, Fig. 4 shows a tertiary structure of an HCV protein complex (PDB ID: 2D3U). In the structure drawing, red balls represent binding residues that are correctly predicted (TP) and green balls represents non-binding residues that are correctly predicted (TN). Non-binding residues that are predicted as binding residues (FP) are represented by blue wireframes, and binding residues that are predicted as non-binding residues (FN) are represented by yellow wireframes.

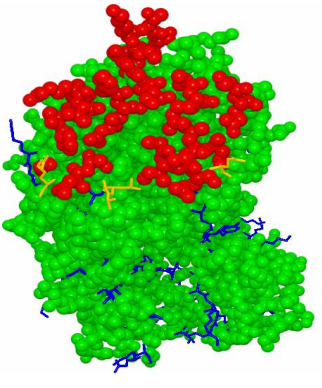


Fig. 2. Prediction of binding residues in an HCV protein complex (PDB ID: 2D3U). Red ball: binding residues that are correctly predicted (TP). Yellow wireframe: binding residues that are predicted as non-binding residues (FN). Green ball: non-binding residues that are predicted correctly (TN). Blue wireframe: non-binding residues that are predicted as binding residues (FP). There are 29 TPs, 427 TNs, 98 FPs, and 4 FNs.

4 Conclusion and Future Work

We proposed a SVM classifier for predicting binding sites in HCV protein complexes using biochemical properties of amino acids. The SVM classifier achieved an average sensitivity of 76.06% and specificity of 75.94%. This SVM classifier outperforms our previous prediction model using a radial basis function neural network (RBFNN).

However, there are several directions to improve the SVM classifier. Firstly, selection of different biochemical properties of an amino acid can improve prediction results. Various biochemical properties of an amino acid exist, and some of them are redundant with each other. Biochemical properties and their features encoded in a feature vector affect the accuracy and efficiency of the SVM classifier. Secondly, using a specialized kernel function for the problem can improve the SVM classifier. In this study, we used general-purpose kernel functions, but a specialized kernel function for this problem can improve the SVM classifier. Although this is a preliminary work, a SVM classifier and the patterns of binding sites found by the classifier would be helpful in better predicting binding sites between HCV and human proteins and in design of new drugs to prevent or treat liver damage caused by HCV-related inflammation.

Acknowledgments. This work was supported by the KOSEF grant (No. R15-2004-033-07002-0) funded by the Korea government (MOST).

References

1. Bogen, S., Saksena, A.K., Arasappan, A., Gu, H., Njoroge, F.G., Girijavallabhan, V., Pichardo, J., Butkiewicz, N., Prongay, A., Madison, V.: Hepatitis C Virus NS3-4A Serine Protease Inhibitors: Use of a P2-P1 Cyclopropyl Alanine Combination for Improved Potency. Bioorg. Med. Chem. Lett. 15, 4515–4519 (2005)
2. Powers, J.P., Piper, D.E., Li, Y., Mayorga, V., Anzola, J., Chen, J.M., Jaen, J.C., Lee, G., Liu, J., Peterson, M.G., Tonn, G.R., Ye, Q., Walker, N.P., Wang, Z.: SAR and Mode of Action of Novel Non-nucleoside Inhibitors of Hepatitis C NS5b RNA Polymerase. J. Med. Chem. 49, 1034–1046 (2006)
3. Hsu, E.C., His, B., Hirota-Tsuchihara, M., Rulan, J., Iorio, C., Sarangi, F., Diao, J., Migliaccio, G., Tyrrell, D.L., Kneteman, N., Richardson, C.D.: Modified Apoptotic Molecule (BID) Reduces Hepatitis C Virus Infection in Mice with Chimeric Human Livers. Nat. Biotechnol. 21, 519–525 (2003)
4. Evans, M.J., Rice, C.M., Goff, S.P.: Phosphorylation of Hepatitis C Virus Nonstructural Protein 5A Modulates Its Protein Interactions and Viral RNA Replication. Proc. Natl. Acad. Sci. 101, 13038–13043 (2004)
5. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E.: The Protein Data Bank. Nucl. Acids. Res. 28, 235–242 (2000)
6. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J.: Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. Nucl. Acids. Res. 25, 3389–3402 (1997)
7. Zhang, G.Z., Nepal, C., Han, K.: Predicting Binding Sites of Hepatitis C Virus Complexes Using Residue Binding Propensity and Sequence Entropy. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4487, pp. 78–85. Springer, Heidelberg (2007)
8. Sander, C., Schneider, R.: Database of Homology-derived Protein Structures and the Structural Meaning of Sequence Alignment. Proteins 9, 56–58 (1991)
9. Homology-derived Secondary Structure of Proteins, http://swift.cmbi.kun.nl/gv/hssp
10. Dodge, C., Schneider, R., Sander, C.: The HSSP Database of Protein Structure-sequence Alignments and Family Profiles. Nucl. Acids. Res. 26, 313–315 (1998)
11. Jones, S., Thornton, J.M.: Analysis of Protein-protein Interaction Sites Using Surface Patches. J. Mol. Biol. 272, 121–132 (1997)

12. Murakami, Y., Jones, S.: SHARP^2: Protein-protein Interaction Predictions Using Patch Analysis. Bioinformatics 22, 1794–1795 (2006)
13. Jones, S., Thornton, J.M.: Prediction of Protein-Protein Interaction Sites Using Patch Analysis. J. Mol. Biol. 272, 133–143 (1997)
14. Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap. Chapman and Hall, London (1993)
15. Georges, D., Masato, K.: Bootstrap Re-sampling for Unbalanced Data in Supervised Learning. Eur. J. Oper. Res. 134, 141–156 (2001)
16. Zweig, M.H., Campbell, G.: Receiver-operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine. Clin. Chem. 39, 561–577 (1993)

Identification of Co-regulated Signature Genes in Pancreas Cancer- A Data Mining Approach

K.R. Seeja[1], M.A. Alam[2], and S.K. Jain[2]

[1] Department of Computer Science
[2] Department of Biotechnology
Hamdard University, NewDelhi, India
seeja@jamiahamdard.ac.in, aalam@jamiahamdard.ac.in,
skjain@jamiahamdard.ac.in

Abstract. Pancreas cancer is one of the most fatal among the cancers. The mortality rate is high due to the lack of tools for proper diagnosis and effective therapeutics. Identification of changes in gene expression in pancreas cancer may lead to the development of novel tools for diagnosis and effective treatment methodology. In this paper we present an association rule mining approach to identify the association between the genes that are either over expressed or under expressed in pancreas cancer compared to normal pancreas. We have used the SAGE data related to pancreas cancer. It is expected that the results will help in developing better treatment methodology for pancreas cancer and also for designing a low cost microarray chip for diagnosing pancreas cancer. The results have been validated in terms of Gene Ontology and the signature genes have been identified that match with published data.

Keywords: Association rule mining, SAGE, pancreas cancer, signature genes, co-regulation.

1 Introduction

In post genomic era, with the availability of the genome databases the focus of research has shifted from sequencing to eliciting knowledge from these databases. The human genome is estimated to have 50,000 to 1,00,000 genes and only 10-20% of these genes are expressed in any cell. Though all cells contain the same genetic material, different genes are expressed in different cell.

Gene expression databases contain information regarding the expression levels of genes in different experimental conditions. Study of these expressed genes is one of the major areas in cancer research. Studies reveal that there is not a single gene behind the initiation and prognosis of cancer, but a group of genes acting together.

Clustering is the technique used earlier for identification of these co-regulated genes. The major drawback of clustering is that it groups genes into single clusters. But actually a gene can be co-expressed with different set of genes. Association rule mining is a better alternative of clustering techniques.

Data mining is a technique used to extract knowledge from large databases. Association rule mining is a data mining technique used in market basket analysis to

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 138 – 145, 2008.
© Springer-Verlag Berlin Heidelberg 2008

find out association between different items purchased by the customer. This result is used for effective decision making. This technique can be used in DNA data analysis to identify the co-regulation of genes.

Let D is a transaction database. Let I={ $i_1,i_2,\ldots i_n$ } is the set of items. Each transaction T is a set of items such that T is a subset of I. Two measures are associated with association rule – support and confidence. Consider an association rule A=> B where A and B are subsets of I and A∩B=ϕ. Support denote the number of transactions in D that contain both A and B.

$$\text{Support}(A=>B)=P(A \cup B).$$

Confidence denotes the number of transactions in D containing A that also contain B. Confidence(A=>B)=P(B|A)=support(A U B)/support(A).

Rules are considered as valid (or strong) rules only if the rules satisfy the user defined minimum support and minimum confidence threshold.

Association rule mining is a two step process. First step is to find all frequent itemsets. Frequent item sets are set of items that occur simultaneously in at least as many transactions as minimum support threshold. Second step is to generate strong association rules from the frequent itemsets.

In this paper, we generated rules of the form

$$\text{gene A}[+]\rightarrow\text{geneB}[+],\text{geneC}[-],\text{geneD}[+]$$

which means that whenever gene A is over expressed in a pancreas cancer tissue, gene B and gene D are also over expressed and gene C is under expressed.

2 The Problem

Identification of the signature genes or marker genes of various cancers plays an important role in the development of therapeutics for cancer. Here the problem is to detect set of genes differentially expressed in cancer tissue. In this paper we are trying to find association between these signature genes. We used a two step procedure.

Step 1: Identification of signature genes.

Here we detected genes which show over expression or under expression in pancreas cancer compared to normal pancreas. This result can be used for designing a microarray chip containing these differentially expressed genes, for diagnosing pancreas cancer.

Step 2: Finding co-regulated signature genes.

Here we generated association rules of the form,

$$\text{gene A}[+]\rightarrow\text{geneB}[+],\text{geneC}[-],\text{ geneD}[+]$$

which means that whenever gene A is over expressed in a pancreas cancer tissue, gene B and gene D are also over expressed and gene C is under expressed. These rules can be used for developing better therapeutics. For example, a medicine which can reduce the expression level of gene A may have a hidden effect of decreasing the expression level of gene B and gene D and increasing the expression level of gene C.

3 Related Work

Association rule mining[1],[2],[3] is one of the commonly used data mining technique for gene expression analysis for identifying co-expressed genes. This technique has already been used to identify synexpression groups[4],[13] from freely available human SAGE data.

Creighton and Hanash[5] used association rule mining for finding association between genes from yeast data set. Tara and Sanjay[6] used this technique for extracting association among genes from microarray data. In another study, the association rule mining is used to associate motif with gene expression.

The traditional association rule mining algorithms [1],[2],[3] are not suitable for mining gene expression databases. They are suitable for database with large number of transactions and small number of items. The structure of gene expression database is a reverse one. It has large number of items (genes) and small number of transactions (tissue samples). Some algorithms were also developed [7],[8],[9] specially for mining gene expression databases.

Identification of differentially expressed genes in disease tissue [12] is another area of research in gene expression data analysis. These genes are called marker genes. This can be done by gene expression profiling. Serial Analysis of Gene Expression (SAGE)[14] is a powerful tool for gene expression profiling. Data derived using the SAGE technology [15] has been used to identify marker genes of a variety of cancers. Some studies [16],[17],[18] have already undergone on SAGE data to identify the signature genes in pancreatic cancer

Another related area of research is to identify differentially co-expressed genes [19],[20],[21],[22]. This approach is based on the concept that cancer related genes may form differential gene-gene co-expression patterns in normal and cancer tissues. This can be used for identifying cancer marker genes.

In this paper, we are trying to find out associations among the marker genes.

4 Data Mining Approach

Data mining is a collection of techniques used to extract knowledge from very large databases. The various steps in data mining are data selection, data pre-processing, data reduction, data transformation, data mining and validation.

4.1 Data Selection

Gene expression profiling techniques can be classified into two- hybridization-based technologies, such as DNA microarrays and sequencing-based approaches like SAGE. We have selected the SAGE data and downloaded from Cancer Genome Anatomy Project website [23]. There were six SAGE libraries of pancreas cancer and two libraries of normal pancreas. The libraries were in the form of text files. Each file contains a set of tags or 10-base sequence and their corresponding count. Each tag represents a gene and the count shows its expression level. The sizes of the libraries are given in Table 1.

Table 1. Number of records in SAGE libraries

Library name	Number of records
CL_PL45	11121
CL_ASPC	10622
B_96_6252	14339
CL_Panc1	10293
CL_CAPAN1	14815
B_91_16113	15793
CS_H126	12360
CS_HY	12392

4.2 Data Preprocessing, Reduction and Preparing the Boolean Table

Since the library sizes are not same, they should be normalized before comparison. For normalizing the data, divide the count value by the total number of tags in the library and multiplied by 300,000, the estimated number of RNAs per cell[4]. Then find the number of distinct tags among all the eight libraries. We got 11745 distinct tags. Then we create a table with the fields- tags, count1, count2, count3, count4, count5, count6, count7, and count8, where tags represent the distinct 10-base sequences found and count1 to count8 represent the count value of that tag in library1 to library 8. We got a table of size 11745 X 8. For each tag find its highest and lowest count value in the normal pancreas libraries. These are the maximum and minimum count of that tag. Replace the count values of the tags in the pancreas cancer libraries with –1 if the tag count is less than its minimum count, with +1 if the tag count is greater than its maximum count and with 0 if the count is less than or equal to maximum count and greater than or equal to minimum count. Thus +1 represents overexpression, -1 represents underexpression and 0 represents no change in expression of the corresponding gene. A sample table is shown in Table 2.

Table 2. A sample table

	Count1	Count2	Count3	Count4	Count5	Count6
Gene1	+1	-1	0	0	+1	0
Gene2	0	0	+1	-1	0	0
Gene3	+1	+1	0	-1	0	+1
Gene4	-1	+1	0	-1	0	-1

Some genes show over expression in some libraries and under expression in some other libraries. Records correspond to such ambiguous tags as well as the records correspond to genes show no change in expression in all the cancer libraries were deleted. Then we deleted the records correspond to genes show difference in expression in only one cancer library also considering it as a sequencing error. Thus the number of tags reduced to 3820. A tag to gene map tool, ACTG [24] is used to get the genes correspond to SAGE tags. Among the 3820 tags, 867 were found under expressed, 2659 were found over expressed and 294 were unknown. Delete these unknown genes

also. Genes correspond to the remaining 3526 tags are the signature genes Thus the size of the table reduced to 3526X6. To create the boolean table suitable for mining, all the +1 and -1 are converted to 1. In the Boolean table 1 represents genes that are differentially expressed in the corresponding cancer libraries and 0 represents genes that are not differentially expressed. The boolean table corresponds to table 2 is shown below.

Table 3. The boolean table corresponds to Table.2

	Count1	Count2	Count3	Count4	Count5	Count6
Gene1	1	1	0	0	1	0
Gene2	0	0	1	1	0	0
Gene3	1	1	0	1	0	1
Gene4	1	1	0	1	0	1

4.3 Association Rule Mining

Data mining technique we used here is the association rule mining**.** Association rule mining is a two step process. First step is to find all frequent itemsets. Frequent itemsets are set of items that occur simultaneously in at least as many transactions as

Table 4. A Sample Rule

{IBRDC2[+]} → {AASS[+], DKFZP564O0463[-], FLJ20160[-], RPS27A[+], SLC31A1[-], LSM1[-], RPL23A[+], SFRS6[-], HSMPP8[-], ANKRD11[-], SNRPB2[-], PREB[-], HNRPM[-], DNMT1[-],DDOST[-], RPL31[-], ITGB4[+], CAP1[+], TGFB1I4[-], PPP4C[+], GLRX[-], PABPN1[+], COX6C[-], FLJ10726[-], DNAJB6[-], CALM1[+], RUTBC1[-], SERPINE1[-], EDF1[-], SPCS2[-], PPP2CB[-], RPL12[-], CORO1C[-], ATP1B3[-], G0S2[-], PP784[-], YARS[-], FOSL1[-], SCRIB[-], GPATC4[-], NSMCE1[-], C21orf25[-], RAB14[-], IFITM3[+], SIAT7D[-], MLLT10[-], HIPK2[-], PRSS15[-], SOLH[-], ATP5J2[-], AKAP1[+], OCA2[+], PUM1[+], PRDX1[+], HLA-A[-], C4orf9[-], EEF2[-], DBN1[-], S100A10[+], S100A16[-], MAP3K12[+], LOC285972[-], GA17[-], SCNM1[+], NFX1[-], SNCA[-],SFRS10[-], RPL17[-], MGC11257[-], RPL36[-], PYCR1[-], KIF1C[-], FREQ[-], AKNA[-], SIAT7F[-], PARK7[-], FPGS[-], MKL1[-], RPL10[-], RPL10[-], SH3GL1[-], ZFYVE27[-], VAMP2[-] TTC19[-], ANGPTL4[+], ARPC1B[+], MGC40499[-], LOC389541[+], ROBO3[-], ATXN2L[-], HNRPH3[-], LOC339229[-], ATXN2[-], SHMT2[-], NIPSNAP1[-], EPC1[-], PPP1R15A[-], CSE1L[-], CLDN4[+], VMP1[+], ARCN1[-], SMURF2[-], CHC1[-], SET7[-], ODC1[-], LGALS3BP[+], TPT1[-], FLJ13456[-], KIAA0664[-], RPS2[-], PDCD1LG1[-], MIR16[-], KIAA1797[-], RPL17[-], RIC8[-], CNTN4[+], CTSC[-], STC1[-], MRLC2[+], RALGDS[-], GFPT2[-], RAP1B[-], RPL23[-], H63[-], NET-5[-], USP7[-], LRPAP1[-], ELAVL1[-], ZNF503[-], TEX10[-], ETV5[-], TGFB1I1[-], POLR2E[-], CPNE2[-], OPRS1[-], NDUFB9[-], FLJ14490[-], EIF2C2[-], ANKRD11[-], SRP14[-], RASAL2[-], SERF2[+], STK4[-], CECR5[-], CDC37[-], FABP5[-], RPL14L[-], H2AFY2[-], ENO1[+], ST13[-], LOC253982[-], RECQL5[-], LPIN2[-], C4orf14[-], TNPO3[-], FHL2[-], PSMD2[-], DAP[+], ATP6V1G1[-], H3F3A[+], XPMC2H[-], LOC130074[-], RNP24[+], CGI-90[-], SPIB[-], ENO2[-], INE1[+], SNRPD3[+], FNBP4[-], RPL30[-], ITGB5[+], RPS9[-], DRPLA[-], ANAPC2[-], PPIA[+], RPL26L1[-], CAPNS1[+], BTBD2[+], ATP2A1[-], SNRPB[-], ZHX3[-], NPM3[-], RAD52[-], GNA11[-], RECQL4[-], MRCL3[+], SMURF2[-], RPL12[-], ZMYND17[-], ETS1[-], TRIP6[-], ANKRD11[-], ATR[-], PRDX2[-], KRT8[+], DNM2[+], EXOSC7[-], YTHDF1[-], KIAA1542[+], JMJD3[-], RABAC1[+], CTPS[-], ARHGAP21[-], CBLB[-], SUPT16H[-], ACTN4[+], TNFRSF10A[-], REA[-], NET-5[-], RPL32[-], CUL4A[-], RPL35[-], ASB8[+], CTDP1[-], AKT1S1[-], RPL11[-], C9orf10[-], SEPHS2[-], HRMT1L2[-], C9orf60[-], KIFC2[-], EEF1A1[-], FLJ20244[-], JUN[-], EWSR1[-], ZF[+], TPT1[-], NAGA[-], ZNF297B[-], RPL13A[-], FZR1[-], ATP6V0E[-], SLITRK6[-], BTF3[+], BTF3[-], WDR6[-], MAP1LC3B[-], GNAI2[+]} support=6/6

minimum support threshold. Since the traditional association rule mining algorithms [1],[2] are not suitable for mining gene expression databases, we used our own algorithm[7] that we have developed specially for mining gene expression databases. 20 maximal frequent itemsets were generated. All frequent itemsets can be derived from these maximal frequent itemsets.

Second step is to generate strong association rules from the frequent itemsets. The major problem here is too many rules. We have limited the rules in two ways.

1. Generate rules with a single gene in RHS.
2. Generate only rules of confidence=1.
   One such rule is shown in table 4.

4.4 Validation

We tried to validate the biological meaning of the marker genes and thereby the biological meaning of the rules in terms of gene ontology. For this we used the GOTerm finder tool [25]. The classifications of over expressed and under expressed genes in terms of gene ontology are shown in figure 1 and figure 2.

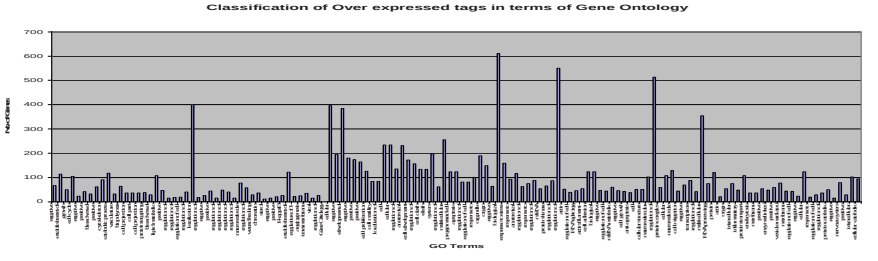


Fig. 1. Classification of over expressed genes

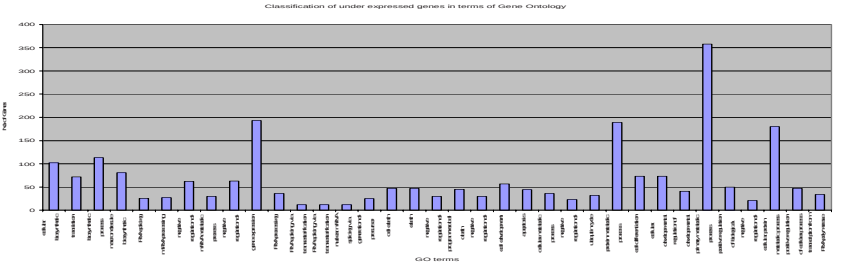


Fig. 2. Classification of under expressed genes

Many marker genes we got are matching with previous reported works [18]. The overexpressed genes we got are mainly classified into regulation of biological/cellular process, localization, biogenesis, cell proliferation, cell death, cell/biological adhesion, multicellular organismal process, response to stress and so on which show their connectivity with tumor. Similarly the underexpressed genes are classified into metabolic process, gene expression, biosynthetic process, developmental process and so on.

5 Conclusion

In this paper we have described a data mining approach for identifying of co-regulated signature genes. We have demonstrated the usefulness of association rule mining using SAGE data related to pancreas cancer. The same approach can be used with microarray data also. We have validated the results in terms of gene ontology and with the help of literature regarding previous reported works. However clinical validation is needed to confirm the results. We hope the results will be helpful in developing better treatment methodology and earlier diagnosis of pancreas cancer. Also the signature genes we got will provide further refinement in the design of a pancreas cancer specific microarray.

References

1. Agrawal, R., Imielinski, T., Swami, A.: Mining Association Rules Between Sets of Items in Large Databases. In: Proceedings of the ACM SIGMOD International Conference on Management of Data, May, pp. 207–216 (1993)
2. Agrawal, R., Srikant, R.: Fast Algorithms for Mining Association Rules in Large Databases. In: Proceedings of the 20th International Conference on Very Large Data Bases (September 1994)
3. Houtsma, M., Swami, A.: Set-Oriented Mining for Association Rules in Relational Data-Bases. In: IEEE 11th International Conference on Data Engineering, pp. 25–33 (1995)
4. Becquet, C., Blachon, S., Jeudy, B., Boulicaut, J.F., Gandrillon, O.: Strong-Association-Rule Mining for Large-Scale Gene-Expression Data Analysis: A Case Study on Human SAGE Data. Genome Biol. 3, 1–16 (2002)
5. Creighton, C., Hanash, S.: Mining gene expression databases for association rules. Bioinformatics 19, 79–86 (2003)
6. Tara, M., Chawla, S.: High–Confidence Rule Mining for Microarray Analysis. IEEE Transactions on Computational Biology and Bioinformatics 4(4), 611–623 (2007)
7. Seeja, K.R., Alam, M.A., Jain, S.K.: A Closed Frequent Itemset Mining Algorithm for Gene Expression Databases. In: Proceedings of the International Conference on Bioinformatics, Computational Biology and Chemoinformatics, July, pp. 30–35 (2008)
8. Mahdi, A., Mohammad, H.S., Gholamhossein, D.: Parallel Mining of Association Rules from Gene Expression Databases. In: Fourth International Conference on Fuzzy systems and Knowledge Discovery (2007)
9. Jiang, X.R., Le, G.: Microarray Gene Expression data association rules Mining Based on JG-Tree. In: proceedings of the 14th International Workshop on Database and Expert Systems Applications (2003)
10. Pedro, C.S., Monica, C., Andres, R., Oswaldo, T., Jose, M.C., Alberto, P.-M.: Integrated analysis of gene expression by association rules discovery. BMC Bioinformatics 7, 54 (2006)
11. Dharmesh, T., Crolina, R., Elizabeth, F.R.: Hypothesis-Driven specialization of Gene Expression Association Rules. In: IEEE International Conference on Bioinformatics and Biomedicine (2007)
12. Balis, R., Jany, B.: Identification of Disease Genes by Expression Profiling. European Respiratory Journal 18, 882–889 (2001)
13. François, R., Céline, R., Sylvain, B., Bruno, C., Olivier, G., Boulicaut, J.F.: Mining Concepts from Large SAGE Gene Expression Matrices. In: Proceedings of 2nd International workshop on Knowledge Discovery in Inductive Databases, September 22, pp. 107–118 (2003)

14. Velculescu, V.E., Zhang, L., Vogelstein, B., Kinzler, K.W.: Serial analysis of gene expression. Science 270, 484–487 (1995)
15. Renu, T., Narendra, T.: Serial Analysis of Gene Expression: Applications in Human Studies. Journal of Biomedicine and Biotechnology 2004(2), 113–120 (2004)
16. Hustinx, S.R., Cao, D., Maitra, A., Sato, N., Martin, S.T., Sudhir, D., Iacobuzio-Donahue, C., Cameron, J., Yeo, C., Kern, S., Goggins, M., Mollenhauer, J., Pandey, A., Hruban, R.H.: Differentially Expressed Genes in Pancreatic Ductal Adenocarcinomas Identified through Serial Analysis of Gene Expression. Cancer Biology and Therapy 3, 1254–1261 (2004)
17. Byungwoo, R.y., Jessa, J., Natalie, J., Parmigiani, B.G., Hollingsworth, M.A., Hruban, R.H., Kern, S.E.: Relationships and Differentially Expressed Genes among Pancreatic Cancers Examined by Large-scale Serial Analysis of Gene Expression. Cancer Research 62, 819–826 (2002)
18. Iacobuzio-Donahue, C., Maitra, A., Shen-Ong, G., et al.: Discovery of Novel Tumor Markers of Pancreatic Cancer using Global Gene Expression Technology. American Journal of Patholology 160, 1239–1249 (2002)
19. Li, H., Krishna, R., Murthy, K.: Significance Analysis and Improved Discovery of Differentially Co-expressed Gene Sets in Microarray Data. In: Sixth IEEE International Conference on Data Mining -Workshops, 2006. ICDM Workshops 2006, December, pp. 196–201 (2006)
20. Michael, W.: CoXpress: differential co-expression in gene expression data. BMC Bioinformatics 7, 509 (2006)
21. Yinglei, L., Baolin, W., Chen, L., Zhao, H.Y.: A statistical method for identifying differential gene–gene co-expression pattern. Bioinformatics 20(17) (2004)
22. Kostka, D., Rainer, S.: Finding disease specific alterations in the co-expression of genes. Bioinformatics 20, Suppl. 1 (2004)
23. SAGE data, `http://cgap.nci.nih.gov/SAGE/SAGELibraryFinder`
24. ACTG tool, `www@retina.med.harvard.edu`
25. GOTerm Finder, `gotools@genomics.princeton.edu`

Neighborhood Rough Set Model Based Gene Selection for Multi-subtype Tumor Classification

Shulin Wang[1,2], Xueling Li[1], and Shanwen Zhang[1]

[1] Hefei Institute of Intelligent Machines, Chinese Academy of Sciences, Heifei, Anhui 230031, China
[2] School of Computer and Communication, Hunan University, Changsha, Hunan 410082, China
jt_slwang@hotmail.com

Abstract. Multi-subtype tumor diagnosis based on gene expression profiles is promising in clinical medicine application. Therefore, a great deal of research on tumor classification based on gene expression profiles has been developed, where various machine learning approaches were applied to constructing the best tumor classification model to improve the classification performance as much as possible. To achieve this goal, extracting features or finding informative genes that have good classification ability is crucial. We propose a novel gene selection approach, which adopts Kruskal-Wallis rank sum test to rank all genes and then apply an algorithm based on neighborhood rough set model to gene reduction to obtain gene subsets with fewer genes and more classification ability. Experiments on a small round blue cell tumor (SRBCT) dataset show that our approach can achieve very high classification accuracy with only three or four genes as evaluated by three classifiers: support vector machines, K-nearest neighbor and neighborhood classifier, respectively.

Keywords: tumor classification; gene expression profiles; support vector machines; neighborhood classifier; K-nearest neighbor.

1 Introduction

Tumor diagnosis is an important application domain of gene expression profiles, which has a promising future in clinical medicine. Therefore, in recent years tumor diagnosis based on gene expression profiles is studied by a number of researchers, the goal of which is to develop an automatic software system that can diagnose tumor for patient and help doctor to analyze the effect of treatment. Especially, the selected informative genes can provide insights into the causes of tumor and are also helpful to the research of tumor treatment drug. At present binary tumor classification is sufficiently researched but it is not to sufficiently satisfy the practical requirements of clinical application. Although multi-subtype tumor classification methods have been researched widely, such as literature [1] and [2], some gene selection methods that are suitable for binary

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 146–158, 2008.
© Springer-Verlag Berlin Heidelberg 2008

classification are not directly applied to multi-class classification, so more approaches to solve multi-class classification should be explored more extensively.

Most of the existing informative gene selection methods are based on gene ranking such as *t*-statistic rank criterion [3, 4] and Fisher's discriminant criterion [5]. Jaeger [6] also compared classification performance with five different test statistics: Fisher, Golub [7], Wilcoxon, TNoM [8], and *t*-test on three tumor datasets. Among these, many gene ranking methods require the dataset following Gaussian distribution. To avoid the assumption of the normality condition, Deng [9] proposed a rank sum test method for informative gene discovery, and the corresponding experiments showed that this method outperforms the previous gene ranking methods. After informative genes are selected out by gene ranking, the next step is to remove redundant genes from the selected gene subset that may be highly relevant as viewed from classification. The general method to solve this problem is wrapper method that combines gene selection with classifier, which is a family of methods, such as combining sequential forward search with three different classifiers [10]. In this paper, we propose a novel gene selection approach to the multi-subtype tumor dataset, which firstly adopts Kruskal-Wallis rank sum test to rank all genes and then applies neighborhood rough set model [11, 12] to gene reduction to obtain many gene subsets with fewer genes and of higher classification ability and finally select the next optimal gene subsets by using classifiers.

2 Methods

2.1 Problem Description

Let $G = \{g_1, \cdots, g_n\}$ be a set of genes and $S = \{s_1, \cdots, s_m\}$ be a set of samples. The corresponding gene expression profiles can be represented as matrix $X = (x_{i,j})$, $1 \le i \le m$, $1 \le j \le n$, where $x_{i,j}$ is the expression level of sample s_i on gene g_i, and usually $n >> m$.

How to select informative gene subset T from gene space $P = 2^G$ is a crucial problem in tumor classification. Suppose $Acc(T)$ denotes the classification ability of gene subset T, which is usually measured by accuracy rate. We hope that the selected informative gene subset T satisfies the following conditions called as the optimal conditions.

$$Acc(T) \ge Acc(G), T \subset G \tag{1}$$

$$\arg\min_{T \in P}(Card(T)) \tag{2}$$

$$\arg\max_{T \in P}(Acc(T)) \tag{3}$$

where $Card(T)$ denotes the cardinal number of T. Note that the optimal gene subset satisfying the optimal conditions is not unique, so the optimal gene set A^* consists of many optimal informative gene subset T^*, that is,

$$A^* = \{T^* \mid T^* \subset G, T^* satisfying (1), (2) and (3)\} \quad (4)$$

2.2 Classification Algorithm and Neighborhood Rough Set Model

Usually, there are two kinds of methods to obtain features for tumor classification: feature extraction and gene selection. However, due to the characteristics of tumor dataset: very high dimensionality and very small sample size, it is difficulty to adopt only one step to extract the best features or to select the smallest informative gene subset to get the highest classification accuracy. We, therefore, design a novel two-step method to implement our goal.

Our algorithm model can be described in Fig. 1 in which step 1 and step 2 achieve the task of informative gene selection or feature extraction. Here informative gene selection is performed by gene reduction based on neighborhood rough set model [11,12] and feature extraction is performed by principal component analysis (PCA) [13]. The main techniques are introduced in detail as follows.

Step 1: *Selecting initial gene set G_{top} according to gene ranking based on the calculated* p*-value of every gene by using Kruskal-Wallis rank sum test.*

Step 2: *Adopting neighborhood rough set model to select out the informative gene set G_o which consists of many gene subsets, or using PCA to extract principal components (PCs) from the gene set G_{top}.*

Step 3: *Evaluating the extracted PCs or selecting the optimal gene set G^* from G_o using classifiers. Usually, the gene subset in G^* has the approximate highest accuracy rate.*

Fig. 1. Pseudocode of algorithm model

Let $NDT =< S, G \cup D, V, f >$ be a neighborhood decision table, where $S = \{s_1, \cdots s_m\}$ is a nonempty sample set called sample space, and $G = \{g_1, \cdots g_n\}$ is a nonempty set of genes also called condition attributes, $D = \{L\}$ is an output variable called decision attribute, V_a is a value domain of attribute $a \in G \cup D$, f is an information function, $f : S \times (G \cup D) \to V$, where $V = \bigcup_{a \in G \cup D} V_a$.

Given $\forall s_i \in S$ and $B \subseteq G$, the neighborhood $\delta_B(s_i)$ of s_i in the subspace B is defined as $\delta_B(s_i) = \{s_j \mid s_j \in S, \Delta_B(s_i, s_j) \le \delta\}$, where δ is a threshold value,

and $\Delta_B(s_i, s_j)$ is a metric function in subspace B. There are three metric functions that are widely used. Let s_1 and s_2 be two samples in *n*-dimensional space $G = \{g_1, \cdots g_n\}$, $f(s, g_i)$ denotes the value of sample s in the *i*-th dimension g_i. Then Minkowsky distance is defined as:

$$\Delta_p(s_1, s_2) = (\sum_{i=1}^{n} \left| f(s_1, g_i) - f(s_2, g_i) \right|^p)^{1/p}$$

Where, (1) if $p = 1$, it is called Manhattan distance Δ_1; (2) if $p = 2$, it is called Euclidean distance Δ_2 ; (3) if $p = \infty$, it is called Chebychev distance.

Given a neighborhood decision table NDT, $X_1, X_2, ..., X_c$ are the sample subsets with decisions 1 to *c*, $\delta_B(x_i)$ is the neighborhood information granules including x_i, and generated by gene subset $B \subset G$, then the lower and upper approximations of the decision D with respect to gene subset B are respectively defined as

$$Lower(D, B) = \bigcup_{i=1}^{c} Lower(X_i, B)$$
$$Upper(D, B) = \bigcup_{i=1}^{c} Upper(X_i, B)$$

where $Lower(X, B) = \{x_i \mid \delta_B(x_i) \subseteq X, x_i \in S\}$ is the lower approximations of the sample subset X with respect to gene subset B, and is also called positive region denoted by $Pos(D, B)$, and $Upper(X, B) = \{x_i \mid \delta_B(x_i) \cap X \neq \phi, x_i \in S\}$ is the upper approximations of the sample subset X with respect to gene subset B.
The decision boundary region of D to B is defined as

$$BN(D, B) = Upper(D, B) - Lower(D, B)$$

The dependency degree of D to B is defined as the ratio of consistent objects $\gamma(D, B) = Card(Pos(D, B)) / Card(S)$, here we define $\gamma(D, \phi) = 0$, and $Card(S)$ denotes the cardinal number of sample set S.
Let $a \in B$, then the significance of a gene is defined as

$$SIG(a, D, B) = \gamma(D, B \cup a) - \gamma(D, B).$$

Step 1: Gene ranking
There are two kinds of rank sum test method: Wilcoxcon rank sum test [14,15] which is a non-parametric alternative to the two-class sample *t*-test and Kruskal-Willis rank sum test [16] which is suitable to multi-class sample testing. Kruskal-Wallis rank sum test is a non-parameter method for testing equality of population medians among classes. We firstly calculate its *H* value for each gene by formula (5).

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} n_i (\overline{r}_{i.} - \overline{r})^2 \tag{5}$$

where n_i is the number of samples in class *i*, r_{ij} is the rank of sample *j* from class *i*, $N = \sum_{i=1}^{k} n_i$ is the total number of all samples, $\overline{r}_{i.} = (\sum_{j=1}^{n_i} r_{ij}) / n_i$, $\overline{r} = (N+1)/2$ is the average of all the r_{ij}. Then we calculate its *p*-value by formula (6).

$$p = \Pr(\chi^2_{k-1} > H) \tag{6}$$

Finally, all genes are ranked in ascending order. There are two methods to select informative genes. One is directly selecting *k* top-ranked genes, and another is selecting the genes whose *p*-values are less than the given significance level threshold α [17].

Step 2: Gene reduction

Although adopting feature reduction based on classical rough set theory to select informative genes is an effective method, its classification accuracy rate is usually not higher compared with other tumor-related gene selection and tumor classification approaches[18]; for gene expression values must be discretized before gene reduction, which leads to information loss in classification. Therefore, the neighborhood rough set model is introduced to tumor classification, which omits the discretization procedure, so no information loss occurs before gene reduction. The forward attribute reduction and neighborhood model (FARNeM) designed based on neighborhood rough set [11] is used as our gene reduction method and shown in Fig. 2.

In FARNeM algorithm, $\gamma(L,B) = Card(Lower(L,B))/Card(S)$ denotes the dependency of subtype label set L with respect to gene subset $B \subseteq G$,where $Lower(L,B)$ denotes the lower approximations of the subtype label set L with respect to gene subset B , and $SIG(a,B,L) = \gamma(L,B \cup a) - \gamma(L,B)$ denotes the significance of gene a with respect to gene subset $B \subseteq G$. The FARNeM algorithm always adds the optimal informative gene into the reduction set in each loop until the dependence does not increase. To better understand the principle of this algorithm need to see literature [11] and [12].

Step 3: optimal gene subset selection based on different classifiers

We respectively adopt three different classifiers to evaluate the selected gene subset to further find the optimal gene subsets that have the highest classification accuracy. (1) Support vector machines (SVM): SVM is a relatively new type of statistic learning theory, originally introduced by Vapnik [19]. It finds out a hyper-plane as the decision surface to maximize the margin of separation between two-class samples. (2) K-Nearest neighbor (K-NN) [20]: K-NN is a most common and non-parametric method. To classify an unknown sample x, K-NN extracts *k* closest vectors from the training set using similarity measures, and makes decision for the label of the unknown sample x using the majority class label of the *k* nearest neighbors. We adopt Euclidean distance to measure the similarity of samples. (3) Neighborhood classifier (NEC)

[11]: similar to K-NN, NEC is also based on the general idea of estimating the class of unknown sample according to its neighbors, but different from K-NN, NEC considers a kind of neighbor within a sufficiently small and near area around the sample, in other words, all training samples surrounding the testing sample take part in the classification decision process.

Algorithm FARNeM
Input: *NDT*=<*S*,*G*,*L*> and δ
// *NDT* is a neighborhood rough decision table, *S* denotes tumor sample set, *G* denotes gene set, and *L* denotes tumor subtype label set. δ is the threshold to control the size of the neighborhood. 2-Norm based neighborhood is used when computing the distance between two samples.
Output: *red* //reduction of gene set *G*
1: $\forall a \in G$, computing neighborhood relation N_a
2: $red = \phi$; //Initial *red* is set to empty set and is the pool to contain the selected informative genes.
3: for each $\forall a_i \in G - red$
Computing $SIG(a_i, red, L) = \gamma(L, red \cup a_i) - \gamma(L, red)$;
// here, define $\gamma(L, \phi) = 0$, the dependency of subtype label set L with respect to empty set is fixed to zero.
4: Selecting the informative gene a_k satisfying
$SIG(a_k, red, L) = \max_i(SIG(a_i, red, L))$;
5: if $SIG(a_k, red, L) > 0$
$red = red \cup a_k$; // Add the optimal gene a_k to the reduction set *red*
go to 3;
else
return *red*;
6: Algorithm end;

Fig. 2. The FARNeM algorithm for gene reduction

2.3 Dataset

From the web site: http://research.nhgri.nih.gov/microarray/Supplement, we download the small round blue cell tumor (SRBCT) dataset which contains 88 samples with 2,308 genes in each sample. In the dataset there are 63 training samples and 25 testing samples, which contain five non tumor-related samples. The dataset details are shown in Table 1. More specifically, the 63 training samples contain 23 Ewing family of tumors (EWS), 20 rhabdomyosarcomas (RMS), 12 neuroblastomas (NB), and eight Burkitt lymphomas (BL). The testing samples contain six EWSs, five RMSs samples, six NB samples, three BL samples. And the five non tumor-related samples are removed in our experiments.

Table 1. SRBCT dataset component description

Class label	Original Dataset	Training set	Testing set
EWS	29	23	6
NB	18	12	6
RMS	25	20	5
BL	11	8	3
Non-SRBCT	5	0	5
Total	88	63	25

2.4 Classification Model Design

To avoid overfitting of classifiers, we design two experiment methods. One is performing an experiment on all 83 samples using 4-fold cross-validation (CV) to evaluate the classification model, which is defined as Method 1. The other is firstly performing an experiment on 63 training samples using CV method to obtain a classification model and then performing an experiment on 20 testing samples using the obtained classification model to gain predictive accuracy rate, which is defined as Method 2, in which the procedure of gene selection is independent of the testing sample set, but in Method 1, the procedure of gene selection is relevant to the testing sample set. Therefore, Method 2 can be more practical than Method 1.

After ranking genes using Kruskal-Wallis rank sum test, how many genes should be selected is a very important problem. To address this problem, the distribution of *p*-value for all genes is shown in Fig. 3, from which we can see that the number of genes with very little *p*-value is great, which lead to the difficulty in determining appropriate number of genes.

To overcome this difficulty, we adopt SVM classifier with Gaussian radial basis function (RBF) kernel $K(x,y)=\exp(-\gamma\|x-y\|^2)$ to evaluate the classification ability of the top-ranked genes. Fig. 4 shows the 4-fold CV accuracy varying with the number of top-ranked genes within 200 top-ranked genes, and it is obvious that the CV accuracies of the 11 to 200 top-ranked genes are all 100% CV accuracy. The heatmap of the 200 top-ranked genes is shown in Fig. 5, in which the consistent expression differentia is obvious among the four tumor subtypes. Therefore, we simply select 200 top-ranked genes as initial informative genes that contain complete classification information. The results also indicate that gene ranking with Kruskal-Wallis rank sum test is very effective.

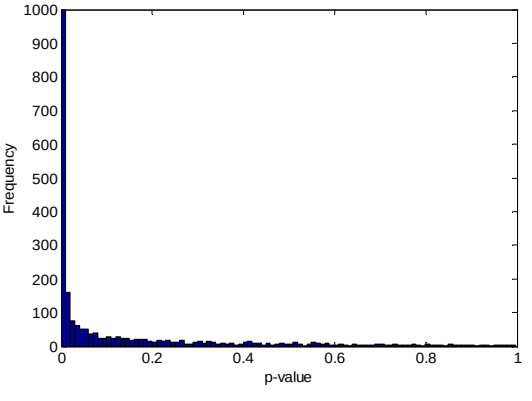


Fig. 3. The distribution of *p*-value for genes

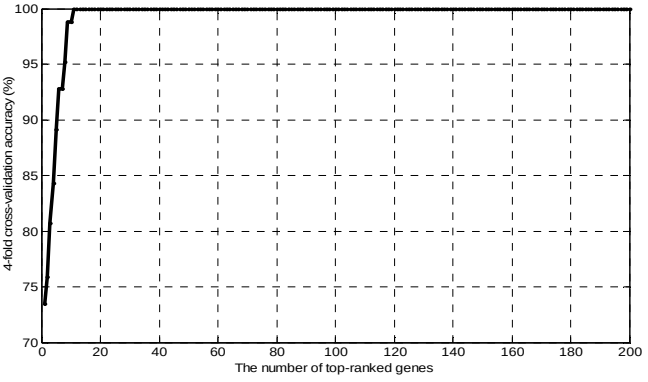


Fig. 4. Classification accuracy varying with the number of top-ranked genes

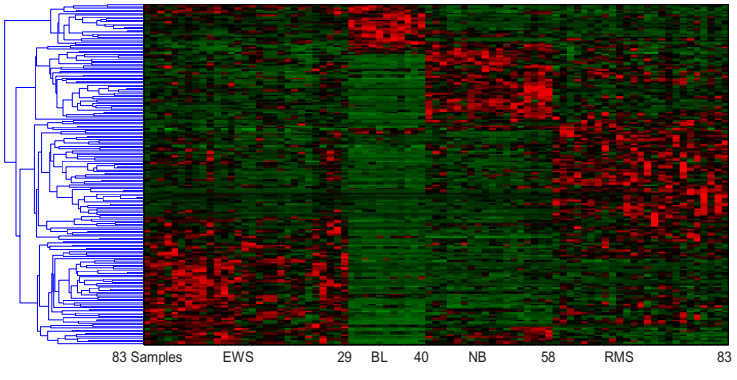


Fig. 5. Gene expression profiles with 200 top-ranked genes

3 Results and Discussion

From the 200 top-ranked gene set, we conduct gene reduction using FARNeM algorithm. This algorithm requires setting the parameter: δ neighborhood. The varying range of δ is [0,1]. One gene subset can be obtained by using FARNeM algorithm corresponding to each δ-value, so 100 gene subsets can be obtained when δ-value varying from 0 to 1 by 0.01. Adopting Method 1, we conduct three classification experiments to evaluate the obtained 100 gene subsets in order to further select the optimal gene subsets with the highest accuracy. Table 2, 3 and 4 show the selected gene subsets and their classification accuracy obtained by using SVM-RBF, K-NN and NEC classifier, respectively, from which we can see that although the performance of NEC classifier is a little lower than that of SVM-RBF, three selected gene subsets are the same as the three obtained by SVM-RBF and NEC respective classifiers. Another advantage is that the number of the selected genes is much lower while

Table 2. The selected gene subsets and their classification using SVM-RBF classifier

No	Selected gene subset	C	γ	δ	CV Acc.
1	{841641,377048,878652,207274, 812105}	200	0.002	0.81	100%
2	**{383188,491565,796258,207274, 812105,295985}**	200	0.0003	0.85	100%
3	**{383188,811000,207274,812105, 770394}**	200	0.0002	0.86	100%
4	**{295985,810057,236282,796258, 866702}**	200	0.004	0.94	100%
5	{841641,812105,1469292,770394,796258, 208699}	200	0.0004	0.98	100%

Table 3. The selected gene subsets and their classification accuracy using K-NN classifier

No.	Selected gene subset	K	δ	CV Acc.
1	**{383188,491565,796258,812105,207274, 295985}**	1	0.85	100%
2	**{383188,811000,812105,207274, 770394}**	4	0.86	100%

Table 4. The selected gene subsets and their classification using NEC classifier

No.	Selected gene subset	Weight (w)	δ	CV Acc.
1	{207274,1435862,812105,377461}	0.02	0.6	98.80 %
2	{841641,383188,207274,812105, 1474684}	0.02	0.79	98.80 %
3	{841641,377048,207274,812105, 1474684}	0.1	0.8	98.80 %
4	**{383188,491565,796258,812105, 207274,295985}**	0.02	0.85	98.80 %
5	**{383188,811000,812105,207204, 770394}**	0.11	0.86	98.80 %
6	**{295985,810057,236282,796258, 866702}**	0.03	0.94	98.80 %

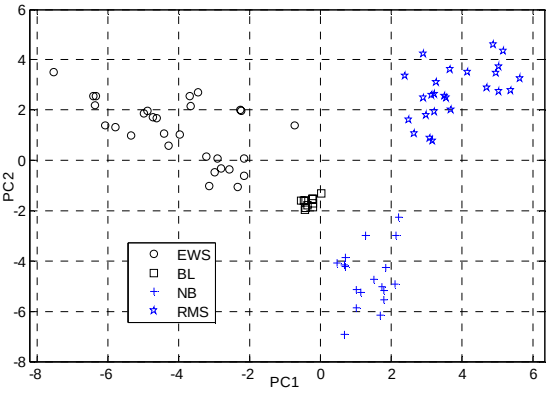


Fig. 6. The scatter plot of two principal components

retaining as high accuracy as possible. Note that SVM-RBF needs two parameters: C andγ, and NEC classifier requires only one parameter: weight w which varies from 0 to 0.6.

In order to visualize the experimental results, we also adopt principal component analysis method to extract principal components (PCs) to be used as the input of classifiers. In our experiments, from the 50 top-ranked genes we extracted only two PCs whose scatter plot is shown in Fig. 6, from which we can see that the boundary between every two tumor subtypes is very clear. Experimental results based on the two

PCs of SVM-RBF, K-NN and NEC classifier, respectively, indicate that 100% 4-fold CV accuracy can be obtained easily.

Similarly, adopting Method 2 we conduct another three classification experiments. Table 5, 6 and 7 show the selected subsets and their classification using SVM-RBF, K-NN and NEC classifier on the 63 training samples, respectively. It is obvious that the selected gene subsets vary with the number of samples. However, gene subset {207274, 1435862, 812105} is always selected by three classifiers, which perhaps means that the three genes are more important than others. We draw scatter plot directly using the three genes as shown in Fig. 7, in which (A) is the scatter plot of 63 training samples and (B) is the scatter plot of 20 testing samples. Moreover, based on the three genes, the classification model is obtained by training with the 63 training samples we use the obtained classification model to predict the other 20 testing samples to obtain predictive accuracy rate. Experimental results indicate that 100% predictive accuracy is gained by using SVM-RBF and K-NN classifier and 95% by using NEC classifier, which means that only one sample is predicted by error.

Table 5. Classification using SVM-RBF classifier on 63 training samples

No.	Selected gene subset	C	γ	δ	CV Acc.
1	{207274,770394,878652,1435862}	200	0.05	0.57	100%
2	{770394,207274,812105,1435862}	200	0.0004	0.63	100%
3	**{207274, 1435862, 812105}**	**200**	**0.003**	**0.64**	**100%**
4	{770394,814526,609663,1434905,1435862}	200	0.001	0.96	100%
5	{770394,814526,435953,866702, 810057}	200	0.0003	0.98	100%

Table 6. Classification using K-NN classifier on 63 training samples

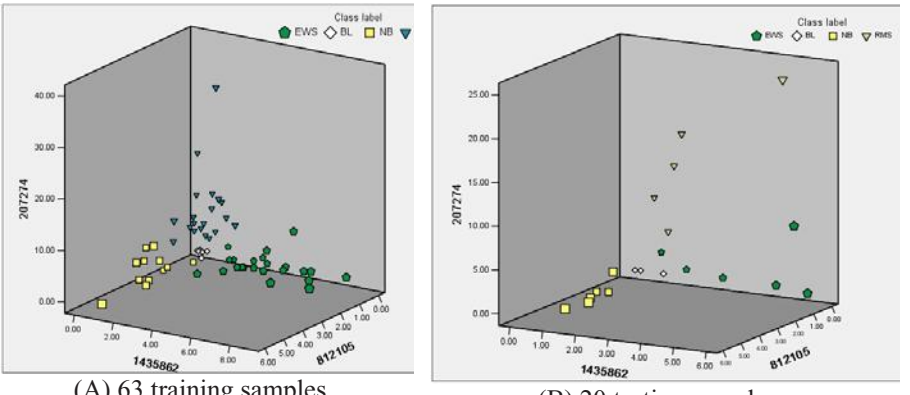
No.	Selected gene subset	K	δ	CV Acc.
1	**{770394,207274,812105,1435862}**	**1**	**0.63**	**100%**
2	{770394,814526,377048,79022, 244618}	1	0.92	100%
3	{770394,814526,435953,866702, 810057}	3	0.98	100%

Table 7. Classification using NEC classifier on 63 training samples

No.	Selected gene subset	Weight (w)	δ	CV Acc.
1	{770394,183337,796258,812105}	0.08	0.5	98.41%
2	{207274,770394,878652,1435862}	0.02	0.57	98.41%
3	{770394,207274,812105,1435862}	0.07	0.63	98.41%
4	**{207274, 1435862, 812105}**	**0.07**	**0.64**	**98.41%**
5	{770394,486110,207274,629896}	0.07	0.77	98.41%
6	{770394,814526,377048,79022, 244618}	0.02	0.92	100%
7	{770394,814526,609663,1434905,1435862}	0.03	0.96	100%
8	{770394,814526,435953,866702, 810057}	0.04	0.98	100%
9	{770394,814526,435953,866702, 812105, 295985}	0.02	0.99	98.41%

To evaluate our methods, we also compare the performance of our methods with others. Table 8 shows the performance comparison with other related work on the same dataset SRBCT. Comparison results show that the number of the selected gene subset obtained by using our method is smaller than the number obtained by adopting

other methods except for the third item in Table 8. And compared with the third item, our method is time-saving, for exhaustive search costs more time. In our method, extracting only 2 PCs can obtained 100% CV accuracy rate, which outperforms the fifth item that needs 3 PCs to obtain 100% accuracy rate.



(A) 63 training samples (B) 20 testing samples

Fig. 7. Comparison of two scatter plots on 63 training samples and 20 testing samples based on three genes {207274, 1435862, 812105}

Table 8. Performance comparison with other related works on SRBCT

No.	Gene selection or feature extraction methods	*Classifier*	#Gene	Acc. %	Ref.
1	The significance for the classification	Artificial neural networks	96	100	[21]
2	GESSES(genetic evolution of subsets of expressed sequences)	K-NN	12 ± 2	100	[22]
3	T-test or Class separability + Exhaustive search	SVM	3	95	[23]
4	The separability-based gene importance ranking	SVM	20	100	[24]
5			3 PCs	100	
6		**SVM**	**3**	**100**	**This**
7	**Kruskal-Wallis rank sum test + FARNeM**	**K-NN**	**4**	**100**	**paper**
8		**NEC**	**5**	**100**	
9	**Kruskal-Wallis rank sum test + PCA**	**SVM or K-NN**	**2 PCs**	**100**	

4 Conclusion and Future Work

Finding informative genes from thousands of genes is a very challenging task, for there is much noisiness and a great number of redundant genes not related to tumor phenotypes. In fact, finding minimum gene subset is equivalent to removing noisy and redundancy from gene expression profiles as much as possible. The advantage of our method is that it can find as few informative genes with discriminative capability as possible. Another advantage is that our method can avoid the assumption of

Gaussian distribution. Experiment on SRBCT dataset shows that the proposed method can select much fewer informative genes that can gain 100% accuracy rate with 4-fold CV by respectively adopting SVM-RBF and K-NN classifier and 98.80% using NEC classifier. If independent test is used as the classification evaluation of the informative genes, 100% predictive accuracy rate is achieved by using the three classifiers, respectively.

Acknowledgments. This work was supported by the grants of the National Science Foundation of China, No. 30700161, the grant of the Guide Project of Innovative Base of Chinese Academy of Sciences (CAS), No.KSCX1-YW-R-30, and the grant of Oversea Outstanding Scholars Fund of CAS, No. 2005-1-18.

References

1. Fu, L.M., Fu-Liu, C.S.: Multi-class Cancer Subtype Classification Based on Gene Expression Signatures with Reliablity Analysis. FEBS Lett. 561, 186–190 (2004)
2. Fung, B.Y.M., Vincent, T.Y.N.: Meta-classification of Multi-type Cancer Gene Expression Data. BIOKDD, 31–39 (2004)
3. Chen, D.C., Liu, Z.Q., Ma, X.B., Hua, D.: Selecting Genes by Test Statistics. J. Biomed. Biotechnol. 2, 132–138 (2005)
4. Furey, T.S., Christianini, N., Duffy, N., Bednarski, D.W., Schummer, M., Hauessler, D.: Support Vector Machine Classification and Validation of Cancer Tissue Samples Using Microarray Expression Data. Bioinform. 16(10), 906–914 (2000)
5. Xiong, M.M., Li, W.J., Zhao, J.Y., Li, J., Boerwinkle, E.: Feature (Gene) Selection in Gene Expression-based Tumor Classification. Mol. Genet. Metab. 73, 239–247 (2001)
6. Jaeger, J., Sengupta, R., Ruzzo, W.L.: Improved Gene Selection for Classification of Microarrays. In: Pacific Symposium on Biocomputing, vol. 8, pp. 53–64 (2003)
7. Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S.: Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science 286, 531–537 (1999)
8. Ben-Dor, A., Bruhn, L., Friedman, N., Nachman, I., Schummer, M., Yakhini, Z.: Tissue Classification with Gene Expression Profiles. J. Comput. Biol. 7(3-4), 559–584 (2000)
9. Deng, L., Ma, J.W., Pei, J.: Rank Sum Method for Related Gene Selection and Its Application to Tumor Diagnosis. Chinese Sci. Bull. 49(15), 1652–1657 (2004)
10. Xiong, M.M., Fang, X.Z., Zhao, J.Y.: Biomarker Identification by Feature Wrappers. Genome Research 11(11), 1878–1887 (2001)
11. Hu, Q.H., Yu, D.R., Xie, Z.X.: Neighborhood Classifiers. Expert Syst. Appl. 34(2), 866–876 (2008)
12. Hu, Q.H., Yu, D.R., Xie, Z.X.: Numerical Attribute Reduction Based on Neighborhood Granulation and Rough Approximation. J. Software 19(3), 640–649 (2008)
13. Jolliffe, I.T.: Principal Component Analysis. Springer, New York (1986)
14. Lehmann, E.L.: Non-parametrics: Statistical Methods Based on Ranks. Holden-Day, San Francisco (1975)
15. Wilcoxon, F.: Individual Comparisons by Ranking Methods. Biometr. 1, 80–83 (1945)
16. Kruskal, W.H., Wallis, W.A.: Use of Ranks in One-criterion Variance Analysis. J. Amer. Statist. Assoc. 47(260), 583–621 (1952)

17. Deng, L., Pei, J., Ma, J.W., Lee, D.L.: A Rank Sum Test Method for Informative Gene discovery. In: KDD 2004, Seattle, USA, pp. 410–419 (2004)
18. Wang, S.L., Chen, H.W., Li, F.R., Zhang, D.X.: Gene Selection with Rough Sets for the Molecular Diagnosing of Tumor Based on Support Vector Machines. In: International Computer Symposium, Taiwan, pp. 1368–1373 (2006)
19. Vapnik, V.N.: Statistical Learning Theory. Springer, New York (1998)
20. Dasarathy, B.: Nearest Neighbor Norms: NN Pattern Classification Techniques. IEEE Computer Society Press, Los Alamitos (1991)
21. Khan, J., Wei, J.S., Ringner, M., Saal, L.H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C.R., Peterson, C., Meltzer, P.S.: Classification and Diagnostic Prediction of Cancers Using Gene Expression Profiling and Artificial Neural Networks. Nature Medicine 7(6), 673–679 (2001)
22. Deutsch, J.M.: Evolutionary Algorithms for Finding Optimal Gene Sets in Microarray Prediction. Bioinform. 19(1), 45–52 (2003)
23. Wang, L.P., Chu, F., Xie, W.: Accurate Cancer Classification Using Expressions of Very Few Genes. IEEE/ACM Trans. Comput. Biol. Bioinform. 4(1), 40–53 (2007)
24. Lee, Y., Lee, C.K.: Classification of Multiple Cancer Types by Multicategory Support Vector Machines Using Gene Expression Data. Bioinform. 19(9), 1132–1139 (2003)

Poisson-Based Self-Organizing Neural Networks for Pattern Discovery

Haiying Wang and Huiru Zheng

School of Computing and Mathematics, University of Ulster, N. Ireland, UK
h.zheng@ulster.ac.uk, hy.wang@ulster.ac.uk

Abstract. Based on unsupervised learning paradigm, self-organizing neural networks have achieved great success in applications of automatic pattern discovery. However, the development of self-organizing neural networks is traditionally based on the assumption that data are governed by a normal distribution. Application of self-organizing neural networks in the areas where data are better modelled by other statistical distributions such as a Poisson distribution has received less attention. Based on the incorporation of the statistical nature of data with a Poisson distribution into a Self-Organizing Map, this paper presents a Poisson-based self-organizing neural network. The proposed network has been tested on two datasets including a real biological example. The results indicate that, in comparison to traditional self-organizing maps, the proposed model offers substantial improvements in pattern discovery in data governed by a Poisson distribution.

Keywords: Poisson statistic, pattern discovery, self-organizing neural network.

1 Introduction

Pattern discovery, as a means of finding useful patterns in data, is a fundamental operation in a data mining process. It involves detecting previously unknown clusters or associated rules in the data. Therefore, the aim of pattern discovery is not merely to recognize patterns but to reveal hidden patterns intrinsic to the tasks investigated. The discovered patterns, which are defined by finding natural groupings of data items based on some similarity metrics or probability density models [1], may have significant applications in diverse domains. In bioinformatics, for example, the discovery of patterns in gene sequences and mass spectrometer data may lead to better understanding of pathways involved in tumor genesis [2]. Based on unsupervised learning paradigm, self-organizing neural networks (SONNs) have demonstrated some unique and interesting features well suitable for pattern discovery.

1.1 SONN: Overview of Principles and Its Applications

SONN represents a family of networks, which have abilities to reveal significant patterns encoded in the data without using a predefined target output. While tuning to statistical regularities of input data, the network optimizes its parameters or undergoes self-organisation [3]. There are two basic learning rules used in this kind of networks: Hebbian learning and competitive learning.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 159–166, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Hebbian learning is based on the modification of synaptic connections between neurones [4]. A SONN based on Hebbian learning may result in several output neurones to be active simultaneously. In the case of competitive learning, however, output neurones of a network compete among themselves for being the one to be activated or fired. At any one time, only a single output neurone is active [3]. Such a feature makes competitive learning highly suitable for identifying statistically significant patterns hidden in the data. Example of SONNs based on competitive learning include Kohonen self-organizing features maps (SOM) [5] and the adaptive resonance theory (ART) models [6].

Among these, the SOM is perhaps the most widely investigated self-organizing networks in the literature used for a range of different purposes including pattern discovery in various domains. In biomedical areas, for example, Golub *et al*. [7] successfully applied SOMs to distinguish types and subtypes of leukaemia based on gene expression patterns. A two-cluster SOM was employed to automatically group 38 leukaemia samples into two classes on the basis of the expression patterns of 50 genes closely correlated with leukaemia disease. Using a mathematical model, lagerholm *et al*. [8] decomposed each electrocardiogram (ECG) beat into a 12-dimensional data vector, and then a SOM was utilized to cluster the encoded ECG beats. The integrated method presented in this study exhibited a very low degree of misclassification.

Most of these applications are based on the assumption that data follow a Gaussian distribution. Therefore traditional distance functions such as Euclidean distance and correlation coefficient have been widely utilized to guide self-organization process. However, in the cases where data are better modelled by a Poisson distribution, traditional distance-based pattern analysis has shown poor performance [9]. It has been demonstrated that without incorporating the statistical nature of a Poisson distribution into the learning process, a SONN based on traditional distance criteria may fail to achieve its goals.

1.2 Objectives of This Study

This paper aims to present a new Poisson-based SOM (PSOM) network, which utilizes the specific statistical nature of data with a Poisson distribution to guide self-organization process. The main purpose of this study is, based on the introduction of a new distance function tailored to a Poisson distribution, to develop a SONN model in order to improve data mining and pattern discovery in the data approximately governed by a Poisson statistics.

The remainder of this paper is organized as follows. Section 2 presents a detailed description of the algorithm of PSOM. Two datasets used to assess the performance of the proposed network are described in Section 3. Results and a comparative analysis are presented in Section 4. The discussion of results and conclusions, together with future research, are given in Section 5.

2 Description of Algorithms

The proposed PSOM follows the basic principle of traditional SOM learning process but incorporates Poisson statistics into two fundamental operations involved in each

input presentation: (1) finding the best matching neurone for each input data and (2) updating the winning neurone and its neighbourhood [5].

Let x_i be the vector representing the i^{th} n-dimensional input data sample, w_j be the associated weight vector of the j^{th} node, and the index k indicate k^{th} value of n-dimensional vector, where x_i and w_j are denoted by the following representation:

$$x_i = [x_{i1}, x_{i2}, \cdots, x_{ik}, \cdots x_{in}]^T, \quad w_j = [w_{j1}, w_{j2}, \cdots, w_{jk}, \cdots w_{jn}]^T \tag{3}$$

During the training process, each input data is assigned to a neurone that has the smallest distance to the input vector. Based on the assumption that each value, $x_{ik}(k = 1, \cdots, n)$, in the input vector follows an independent Poisson distribution, this study adopts the joint probability-based distance function to measure the distance between i^{th} input and j^{th} neurone:

$$d(i, j) = 1 - \prod_{k=1}^{n} (\exp(-\lambda_{ik,j}) \times \lambda_{ik,j}^{x_{ik}} / x_{ik}!) \tag{4}$$

$$d(i, c) = \min(d(i, j)), \quad \forall j \tag{5}$$

Where $\lambda_{ik,j}$ represents the expected value of x_{ik} given the weight vector of j^{th} neuron, 'min' represents the minimum operator and the index c indicates the winning neurone. Given the fact that after a learning process, the weight vector associated with each neurone uniquely represents the cluster profile of all the samples assigned to this neurone, and we are interested in grouping all the samples with similar relative data profiles, $\lambda_{ik,j}$ can be calculated [10] as:

$$\lambda_{ik,j} = \frac{w_{jk}}{\sum_{k=0}^{n} w_{jk}} \times \sum_{k=0}^{n} x_{ik} \tag{6}$$

After finding the winning neurone for each input sample, all the weight vectors within the neighbourhood of a winning neurone need to be updated according to the given i^{th} input sample. Apparently, the traditional weight strategy based on absolute values is not suitable in our case. Given that the goal of PSOM is to group all the data items with similar relative values, we utilize the following weight adaptation strategy:

$$w_{jk}(t+1) - w_{j,k}(t) = \begin{cases} \alpha(t) \times \left(\frac{x_{ik}(t)}{\sum_{i=1}^{n} x_{ik}(t)} - \frac{w_{jk}(t)}{\sum_{k=1}^{n} w_{jk}(t)}\right) \times \sum_{k=1}^{n} w_{jk}(t), & j \in N_c(t) \\ 0, & otherwise \end{cases} \tag{7}$$

where $w_{jk}(t+1)$ and $w_{jk}(t)$ are the weight values of the k^{th} value of the j^{th} neurone before and after the adaptation at iteration t. $N_{c,t}$ and α_t represent the neighbourhood of the winning neurone c and learning rate at iteration t respectively. The reader is referred to [5] for a more detailed description of selection criteria for α and N_c. In this paper, the initial neighbourhood (N_0), the maximum number of learning epochs (*NLE*), and the initial learning rate (α_0), were set to 3, 2000, and 0.05 respectively.

The PSOM was implemented within the software development framework provided by the open-source platform, *TIGR MeV* [11].

3 Datasets Studied

The proposed PSOM model is tested on two datasets: a synthetic dataset and a regulatory sequence dataset.The synthetic data were constructed based on a study published by Cai *et al*. [9]. A total of 200 synthetic samples, each representing by five simulated values at five time points: T1, T2, T3, T4, and T5, were analyzed. All the simulated values were generated independently using Poisson distributions. Based on relative simulated levels across five time points, these 200 samples were divided into four groups: A, B, C, and D with 30, 40, 60, and 70 samples respectively.

The regulatory sequence dataset was obtained from a study published by van Helden [12]. Regulatory regions, served as binding sites for transcription factors, can be characterized on the basis of counts of pattern occurrences. It has been suggested that the probability of finding an occurrence of specified regulatory motifs at any position within a gene sequence can be modelled by a Poisson distribution. This dataset consists of 64 genes, of which occurrences of 44 resulting regulatory patterns in the upstream region of each gene were counted 13. On the basis of the biological pathway involved, these 64 genes were categorized to three groups: 31 genes belong to NIT family involved in nitrogen metabolism, 20 belong to MET family involved in methionine biosynthesis, and 13 belong to PHO family involved in phosphate utilization. The 44 regulatory patterns were extracted from the upstream regions of these genes by using the computational analysis of oligonucleotide frequencies program *oligo-analysis* [14] and *dyad-analysis* [15]. The reader is referred to [12] for a full description of the generation of this dataset.

4 Results

4.1 Analysis of Synthetic Dataset

We firstly implemented a comparative analysis by applying the SOM with traditional distance metrics and the PSOM to the synthetic dataset, as shown in Fig. 1. A representative 3 x 2 map of the PSOM for synthetic dataset is illustrated in Fig. 1(a). The PSOM firstly clustered 200 samples into 2 neurones (Neurones (1,1) and (3,2)), which are well-separated by 4 empty neurones. While all the 100 samples from Groups A and D fall into Neurone (1, 1), the samples belonging to Groups B and C are assigned to Neurone (3, 2), suggesting that this dataset exist two main classes. Further analysis on these two neurones using a 1x2 PSOM based on *one-neurone-take-one-cluster* strategy indicates the existence of two subclusters for Neurones (1,1) and (3,2). All the 200 samples are correctly clustered into 4 groups. On the contrary, clustering analysis with the SOM using traditional distance functions, as shown in Fig. 1(b) and (c), fails to achieve its clustering goals. For example, in Fig. 1(b), 60 Group C samples are assigned to two neurones: Neurones (1, 2) and (3, 2), which are separated by Neurone (2, 2). Although like the PSOM, all the samples from Group A and D fall into Neurone (1, 1), further clustering analysis on this neurone reveals that the SOM

with Euclidean distance is unable to differentiate these two groups. The clusters generated by the SOM with Pearson correlation, as shown in Fig. 1(c), are less meaningful. Seventy samples from Group C are dispersed into 4 neurones. Interpreting such a map may prove to be rather difficult. These results highlight the advantages of the PSOM over the traditional SOM when dealing with Poisson-distributed data.

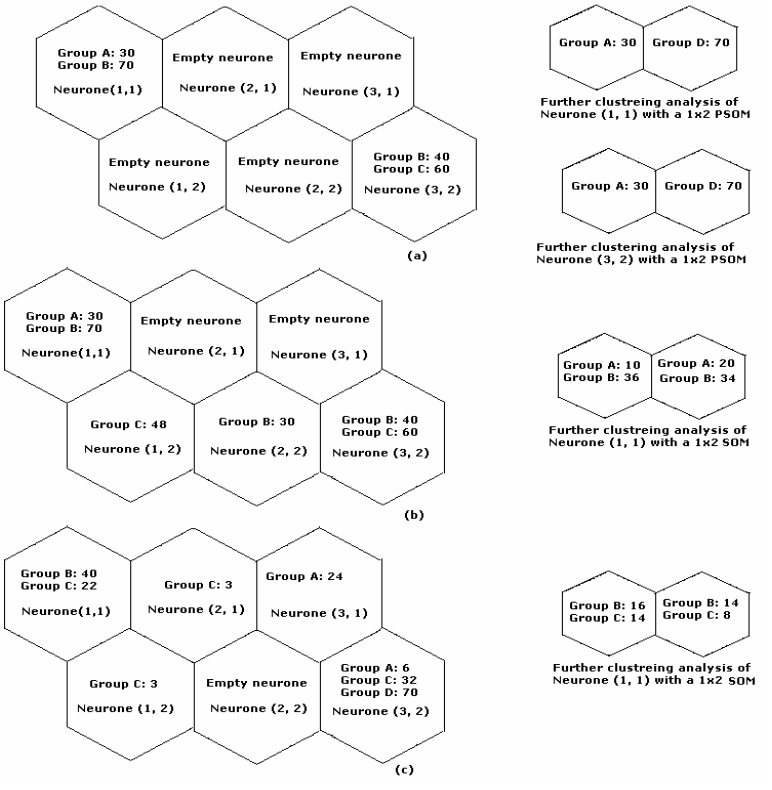


Fig. 1. Clustering analysis for the synthetic dataset with (a) PSOM; (b) the SOM with Euclidean distance; and (c) the SOM with Pearson correlation. The hexagons represent the corresponding map neurones. Each neuron shows the samples clustered.

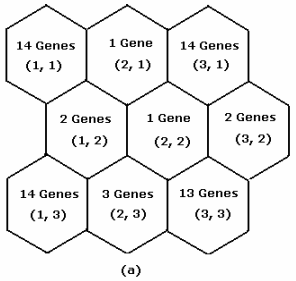
4.2 Analysis of Regulatory Sequence Data

To further assess its performance, we applied the PSOM to a real biological dataset: regulatory sequence data, which are better approximated by a Poisson distribution. The outcomes of a 3×3 map for 64 genes with the PSOM is depicted in Fig. 2(a) . An analysis with regard to the distribution of three gene families over each neurone (Fig. 2(b)) indicates that each neurone is highly over-represented by certain gene family. For instance, all the 14 gene sequences assigned to Neurone (1, 1) belongs to MET family involved in methionine biosynthesis process. While NIT genes are highly over-represented in Neurones (1,3) and (3, 3), genes belonging to PHO family are significantly enriched in Neurone (3,1). The level of gene family enrichment for

each neurone can be quantitatively calculated by the *hypergeometric distribution* probability [15] as follows:

$$p = 1 - \sum_{i=0}^{k-1} \frac{\binom{K}{i}\binom{N-K}{n-i}}{\binom{N}{n}} \tag{9}$$

where K is the number of gene sequences assigned to a neurone, k is the number of family members found in a given neurone, N is the total number of gene sequences and n is the number of sequences belonging to a specific family in the whole dataset. If this probability is sufficiently low for a given gene family, one may say that such a family is significantly represented in the cluster. The p-value, calculated using Equation (9), for observing the dominant gene class in each neurone is shown in Fig. 2(c). Obviously, such a low probability strongly indicates that the distributions of three gene families over each neurone are most unlikely to occur by chance.



(a)

Neurone	MET	NIT	PHO
(1, 1)	14	0	0
(1, 2)	2	0	0
(1, 3)	0	14	0
(2, 1)	1	0	0
(2, 2)	0	1	0
(2, 3)	0	3	0
(3, 1)	3	0	11
(3, 2)	0	2	0
(3, 3)	0	11	2
Total	20	31	13

(b)

Neurone	Dominant class	p value
(1, 1)	MET	$<10^{-10}$
(1, 2)	MET	$<10^{-10}$
(1, 3)	NIT	$<10^{-10}$
(2, 1)	MET	$<10^{-10}$
(2, 2)	NIT	$<10^{-10}$
(2, 3)	NIT	$<10^{-10}$
(3, 1)	PHO	$<10^{-9}$
(3, 2)	NIT	$<10^{-10}$
(3, 3)	NIT	< 0.001

(c)

Fig. 2. (a) A 3x3 SOM map for 64 genes with the SOM. The number shown on each neurone stands for the total number of genes that fall into this neurone. (b) The distribution of genes over each neurone. (c) The p values for observing the dominant gene class in each neurone.

5 Discussion and Conclusions

Being able to reveal significant patterns encoded in the data without a predefined target or external teacher, SONNs have become a powerful tool to support pattern discovery in diverse areas. However, the development of SONN traditionally places emphasis on data domains which follow a Gaussian distribution. Applications of SONNs in other areas that are better modeled by different statistical models have received relatively little attention. Based on the incorporation of statistical properties of a Poisson distribution into the learning process, this paper presents a Poisson-based SONN, PSOM. The proposed model has been tested on two datasets, including a real biological example. The obtained results indicated that, in comparison to traditional

SOM, the PSOM offers several substantial improvements in pattern discovery in data governed by a Poisson distribution. By visualizing the resulting map generated by the PSOM, underlying data structure may be readily understood and significant patterns hidden in the data could be easily identified.

Nevertheless, the development of the PSOM is based on the same principle of traditional SOM. It shows some same drawbacks found in the SOM algorithm. For example, it doesn't represent cluster boundaries explicitly. The network structure and the number of neurons had to be specified in advance. The learning process is governed by several learning parameters. Currently, the selection of learning parameters is based on trial-and-error strategy. Combining with other machine learning techniques such as Genetic Algorithm (GA) to automatically determine the optimal learning parameters would be part of our future work.

More recently, Wang *et al.* [10] proposed a SONN model which attempts to incorporate Poisson statistics into SOM based on the utilization of Chi-Square statistics to measure the deviation between actual values and the expected values. By using joint probability function, this paper directly incorporates Poisson statistics into the learning process. It has been shown that by directly incorporating joint probability based distance into the learning process, a better result can be achieved [9]. Zheng *et al.* [17] introduced a Poisson-based self-adaptive neural network (PGSOM). While successfully addressing some of the drawbacks exhibit by SOM, the performance of PGSOM relies on some predetermined parameters such as *spread factor*. This may make the model more susceptible to the variations of learning parameters. It has been shown that SOM-based networks are more robust models [18]. Nevertheless, the comparison of PSOM with these models deserves further investigation.

The Poisson distribution has been widely used as a mathematical model to describe a wide range of phenomena in science, such as transcription-factor binding sites and serial analysis of gene expression data in bioinformatics. The methodologies described in this paper have the potential to contribute to the development of data mining techniques to support knowledge discovery in these areas.

References

1. Fayyad, U., Piatetsky-Shapiro, G., Matheus, C.: The KDD Process for Extracting Useful Knowledge from Volumes of Data. Communications of the ACM 39(11), 27–41 (1996)
2. Polyak, K., Riggins, G.J.: Gene Discovery Using the Serial Analysis of Gene Expression Technique: Implications for Cancer Research. Journal of Clinical Oncology 19(11), 2948–2958 (2001)
3. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice Hall, New Jersey (1994)
4. Hebb, D.O.: The organisation of Behaviour: A Neuropsychological Theory. Wiley, New York (1949)
5. Kohonen, T.: Self-Organising Maps. Springer, Heidelberg (1995)
6. Grossberg, S.: Adaptive Pattern Classification and Universal Recoding: I. Parallel Development and Coding of Neural Feature Detectors. Biological Cybernetics 23, 121–134 (1976)

7. Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gassenbeck, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., Lander, E.S.: Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science 286, 531–537 (1999)
8. Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L., Sörnmo, L.: Clustering ECG Complexes Using Hermite Functions and Self-organising Maps. IEEE Trans. Biomedical Engineering 47(7), 838–848 (2000)
9. Cai, L., Huang, H., Blackshaw, S., Liu, J.S., Cepko, C., Wong, W.: Clustering Analysis of SAGE Aata: A Poisson Approach. Genome Biology 5(51) (2004)
10. Wang, H., Zheng, H., Azuaje, F.: Poisson-Based Self-Organizing Feature Maps and Hierarchical Clustering for Serial Analysis of Gene Expression Data. IEEE/ACM Transactions on Computational Biology and Bioinformatics 4(2), 163–175 (2007)
11. Saeed, I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., Quackenbush, J.: TM4: a Free, Opensource System for Microarray Data Management and Analysis. BioTechniques 34(2), 374–378 (2003)
12. Van Helden, J.: Metrics for Comparing Regulatory Sequences on the Basis of Pattern Counts. Bioinformatics 20(3), 399–406 (2004)
13. Van Helden, J., Andre, B., Collado-Vides, J.: A Web Site for the Computational Analysis of Yeast Regulatory Sequences. Yeast 16, 177–187 (2000)
14. Van Helden, J., Andre, B., Collado-Vides, J.: Extracting Regulatory Sites from the Upstream Region of Yeast Genes by Computational Analysis of Oligonucleotide Frequencies. J. Mol. Biol. 281, 827–842 (1998)
15. Van Helden, J., Andre, B., Collado-Vides, J.: Discovering Regulatory Elements in Non-Coding Sequences by Analysis of Spaced Dyads. Nucleic Acids Res. 28, 1808–1818 (2000)
16. Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J., Church, G.M.: Systematic Determination of Genetic Network Architecture. Nat. Genet. 22, 281–285 (1999)
17. Zheng, H., Wang, H.Y., Azuaje, F.: Improving Pattern Discovery and Visualization of SAGE Data through Poisson-Based Self-Adaptive Neural Networks. IEEE Transaction on Information Technology in Biomedicine (in press, 2008)
18. Köhle, M., Merkl, D.: Visualising Similarities in High Dimensional Input Spaces with a Growing and Splitting Neural Network. In: Vorbrüggen, J.C., von Seelen, W., Sendhoff, B. (eds.) ICANN 1996. LNCS, vol. 1112, pp. 581–586. Springer, Heidelberg (1996)

Proteome-Wide Analysis of Amino Acid Absence in Composition and Plasticity

Yuzhong Zhao[1], Yun Xu[1], Zhihao Wang[1], Changjiang Jin[2], Xinjiao Gao[2], Yu Xue[2,*] and Xuebiao Yao[2,*]

[1] Hefei National High-performance Computing Center, Department of Computer Science & Technology, University of Science & Technology of China, Hefei 230027, China
[2] Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science & Technology of China, Hefei, Anhui 230027, China
{xueyu,yaoxb}@ustc.edu.cn

Abstract. Each post-translational modification (PTM) of proteins only modifies one or several limited types of amino acids to ensure regulatory fidelity. Thus proteins without one or several types of amino acids (amino acid absence, AAA) might prohibit some PTMs. We performed a proteome-wide analysis to compute AAA proteins in six eukaryotic organisms. Although most of AAA proteins are generated by random during evolution, there are 1,220 (~10%) AAA proteins conserved in human and mouse, implying the amino acid absence might be important for a small proportion of proteins as an evolutionary mechanism. In addition, five of 25 K-deficient proteins are static keratin-associated proteins, which interact with keratins and play important roles in animal hair organization.

Keywords: post-translational modification, amino acid absence, Keratin-associated proteins, ubiquitination.

1 Introduction

Each protein has a unique and ordered sequence genetically defined by the coding sequence of the DNA, with different numbers and types of twenty amino acids. And the distinct order of a protein primary sequence determines its specific 3D structure and biological function in vivo.

Besides to build 3D structures of proteins as structural elements, most of the twenty amino acids also afford distinctly post-translational modifications (PTMs). Proteins modified by specific PTMs play important roles in regulating the functions and dynamics of proteins, and are involved in numerous cellular processes [1]. One protein could be modified on a single site or multi-sites by one or several competitive PTMs [2]. And even for the same protein sequence, isoforms contain different PTMs will carry out distinct functions. Thus, PTMs of proteins enhance the complexity and molecular diversity of a proteome [3, 4]. To date, there are more than 350 types of

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 167–178, 2008.
© Springer-Verlag Berlin Heidelberg 2008

PTMs discovered (http://abrf.org/index.cfm/dm.home). And totally fifteen amino acids could be modified except five hydrophobic residues (L, I, V, A, and F) [3].

Identification of PTMs sites in proteins is fundamental for understanding regulatory mechanisms of protein dynamics at molecular level. However, only several types of PTMs were well characterized by experimental or computational approaches, e.g., phosphorylation [5-7], glycosylation [8], sumoylation [9], S-palmitoylation [10], acetylation [11], methylation [12], and ubiquitination [13], etc. Most of PTMs are carried out by specific enzymes or spontaneously in vivo. Each PTM only modifies one or several kinds of residues specifically. For example, phosphorylation performed by serine/threonine (S/T) Kinases or Tyrosine (Y) Kinases could modify specific S/T or Y residues of a protein in eukaryotes [5-7]. And O-glycosylation usually occurs on S/T residues [8], while palmitoylation recognizes cysteine (C) residues [10]. Sumoylation, acetylation and ubiquitination could take place on lysines (K) [9, 11, 13]. And methylation commonly modifies arginines and lysines [12].

Since each PTM recognize distinct types of amino acids, in theory, proteins without a specific kind of amino acids may prohibit some PTMs. For example, proteins without K will not be modified by sumoylation and ubiquitination, etc. And proteins lack Y residue will prevent tyrosine phosphorylation. Also, proteins without C might not be modified by S-palmitoylation. Then an interesting question is emerging: How many proteins with at least one type of amino acid absence (AAA) exist in eukaryotes?

To address the problem, we took a proteome-wide analysis of amino acid absence in protein sequences of six eukaryotic organisms, including *S. cerevisiae* (Budding yeast), *S. pombe* (Fission Yeast), *C. elegans* (Nematode), *D. melanogaster* (Fruit Fly), *M. musculus* (Mouse) and *H. sapiens* (human). The phenomena of amino acid absence are much ubiquitous in eukaryotes. There are about 18.5% ~ 25.5% of proteins with at least one type of amino acid absence. In addition, we noticed that many protein sequences still remain fragment status. To avoid the bias, we removed all fragment sequences from six eukaryotic proteomes and re-calculated AAA proteins. Again, there is still about 14.1% ~ 23.4% of six proteomes composed of AAA proteins. The AAA Proteins tend to be shorter sequences with average lengths from 129.2 aa ~ 197.6 aa, while average lengths of total proteins in six organism fluctuate from 401.1 aa to 503.5 aa. Also, amino acid abundances and distributions were calculated and proposed to be negatively correlated with amino acid absences. Furthermore, we detected 1220 AAA proteins conserved in mouse and human. Specifically, there are 25 conserved AAA proteins without Lysine (K) residues. These proteins with diverse functions might not be ubiquitinated and degraded in vivo. Interestingly, five of 25 K-absence proteins are Keratin-associated proteins (KRA24, KRA42, KRA131, KRA71 and KRA81), which interact with keratins to organize the animal hair.

2 Materials and Methods

2.1 Data Preparation of Protein Sequences for Six Organisms

In this work, we focused on analyzing AAA proteins in eukaryotes. Six eukaryotic proteomes were chosen, including *S. cerevisiae* (Budding yeast), *S. pombe* (Fission

Yeast), *C. elegans* (Nematode), *D. melanogaster* (Fruit Fly), *M. musculus* (Mouse) and *H. sapiens* (human).

Protein sequences were retrieved from Sequence Retrieval System (SRS5) at ExPASy website (http://www.expasy.ch/srs5/). The taxonomic code (TaxID) of each organism, e.g., 4932 of *S. cerevisiae* and 9606 of *H. sapiens*, was employed to query protein sequences from Swiss-Prot & TrEMBL database. Totally, there were 68,504, 63,966, 29,117, 23,263, 5,243 and 7,400 proteins obtained from *S. cerevisiae*, *S. pombe*, *C. elegans*, *D. melanogaster*, *M. musculus* and *H. sapiens*, respectively.

Previously, we found that many proteins remained to be fragment status, and more than 30% of all human proteins were fragment with short sequence [14]. Intuitively, these short, fragment sequences might frequently lack of one or several types amino acids. To avoid the bias, we used the keyword of "fragment" with each TaxID to query fragment proteins in six organisms, respectively. Then fragment sequences were removed, and we re-calculated AAA proteins again. The information of data preparation is shown in Table 1.

Table 1. Fragment protein sequences of several organisms in Swiss-Prot & TrEMBL database. The TaxID of each organism was used to retrieve all protein sequences of these organisms, while the keyword of "fragment" plus the TaxID to retrieve all fragment sequences separately. And the proteome of R.norvegicus (rat) is unintegrated to be removed in our study.

Swiss-Prot & TrEMBL	TaxID	Fragment	Total	Percentile
H.sapiens	9606	23028	68504	33.62%
M.musculus	10090	15662	63966	24.48%
R.norvegicus	10116	3043	14705	20.69%
D.melanogaster	7227	3327	29117	11.43%
C.elegans	6239	508	23263	2.18%
S.pombe	4896	275	5243	5.25%
S.cerevisiae	4932	326	7400	4.41%

2.2 Computational Detection of AAA Proteins

We considered a protein as an AAA protein if at least one type of twenty amino acids was absent from its sequence. By this way, twenty types of AAA proteins in six organisms were computed by PERL language, separately (Table 2).

To avoid the bias, we retained full length proteins to remove all fragment sequences. Again, AAA proteins in fragment and full-length protein sequences were also calculated, respectively. For comparison, percentiles of AAA proteins in all protein sequences and full-length proteins are shown in Fig. 1A. Also, percentiles of AAA proteins in fragment protein sequences are diagramed in Fig.1B.

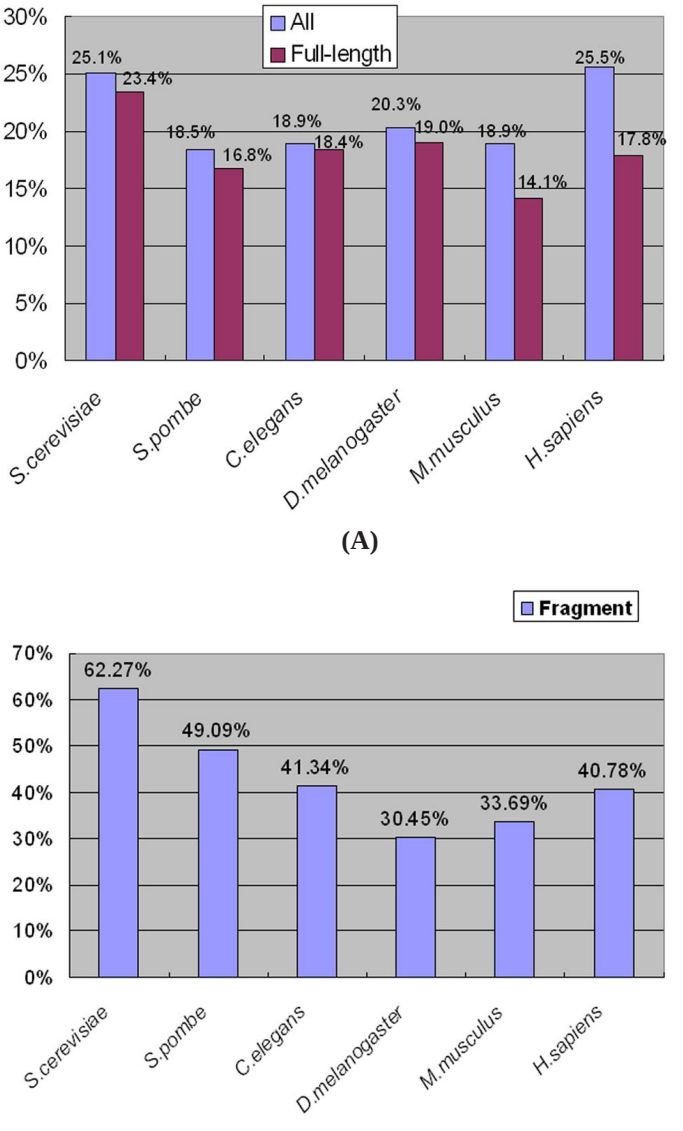


Fig. 1. Percentiles of AAA proteins in six organisms. (A) AAA proteins in all protein sequences vs. full-length proteins; (B) AAA proteins in fragment sequences.

2.3 Identification of the AAA Proteins Conserved between Human and Mouse

To investigate whether and how many AAA proteins are conserved, we chose human and mouse to survey orthologue pairs between two organisms. We downloaded the

Table 2. Twenty types of AAA proteins of all proteins from six organisms. Due to the page limitation, only the numbers of AAA proteins in six organims are shown. SC, S. cerevisiae; SP, S. pombe; CE, C. elegans; DM, D. melanogaster; MM, M. musculus; HS, H. sapiens.

Absent aa	SC	SP	CE	DM	MM	HS
A	98	15	63	90	419	913
C	761	461	1413	1669	3248	5490
D	184	32	193	433	1101	2403
E	145	26	163	353	673	1514
F	70	43	158	450	1262	2821
G	115	31	99	150	437	1003
H	363	130	928	876	2261	4012
I	62	14	100	210	1317	3056
K	55	12	125	265	926	2100
L	14	5	28	56	269	497
M	36	28	83	218	1368	3285
N	76	14	121	334	1515	3274
P	93	37	151	245	624	1268
Q	215	42	182	321	885	1829
R	107	21	153	182	587	1296
S	24	3	37	50	244	679
T	47	8	115	168	503	1314
V	48	10	43	160	561	1080
W	1014	574	2769	3393	6587	8122
Y	176	55	504	977	2629	4788
All AAA proteins	1859	968	4407	5920	12093	17497
Total proteins	7400	5243	23263	29117	63966	68504

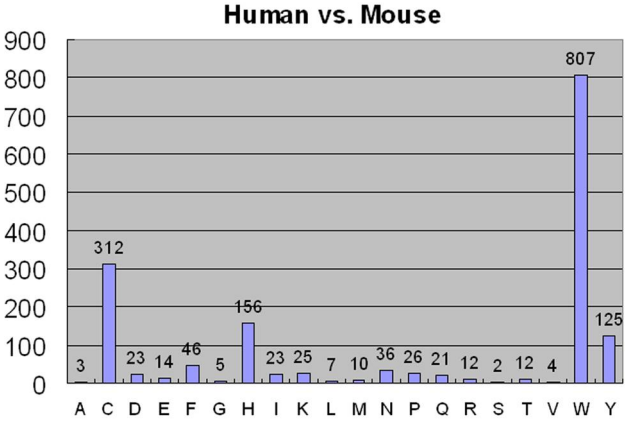


Fig. 2. Totally, there are 1,220 AAA proteins discovered to be conserved in human and mouse

Stand-alone InParanoid Program (ver. 1.0) (http://inparanoid.sbc.su.se/) [15] and computed orthologue pairs between *H. sapiens* and *M. musculus*, with default parameters. Totally, there were 17,733 pairs of orthologs calculated. Then 1,220 AAA proteins were found to be conserved between human and mouse. The information of twenty kinds of 1,220 conserved AAA proteins is shown in Fig. 2.

3 Results and Discussion

3.1 The AAA Proteins Are Ubiquitous in Eukaryotes and Tend to Be Shorter Sequences

In this work, we concerned a series of intriguing problems: Are there any proteins lack of at least one specific type of amino acids (amino acid absence, AAA)? And how many AAA proteins exist in eukaryotes? Are these AAA proteins conserved? And what about their functions are?

To address these questions, we performed a proteome-wide analysis of amino acid absence in protein sequences of six eukaryotic organisms, including *S. cerevisiae* (budding yeast), *S. pombe* (fission Yeast), *C. elegans* (nematode), *D. melanogaster* (fruit Fly), *M. musculus* (mouse) and *H. sapiens* (human). A protein was computed as an AAA protein, if at least one type of twenty amino acids was absent from its sequence. By this simple rule, we calculated twenty types of AAA proteins in six organisms, separately (Table 2).

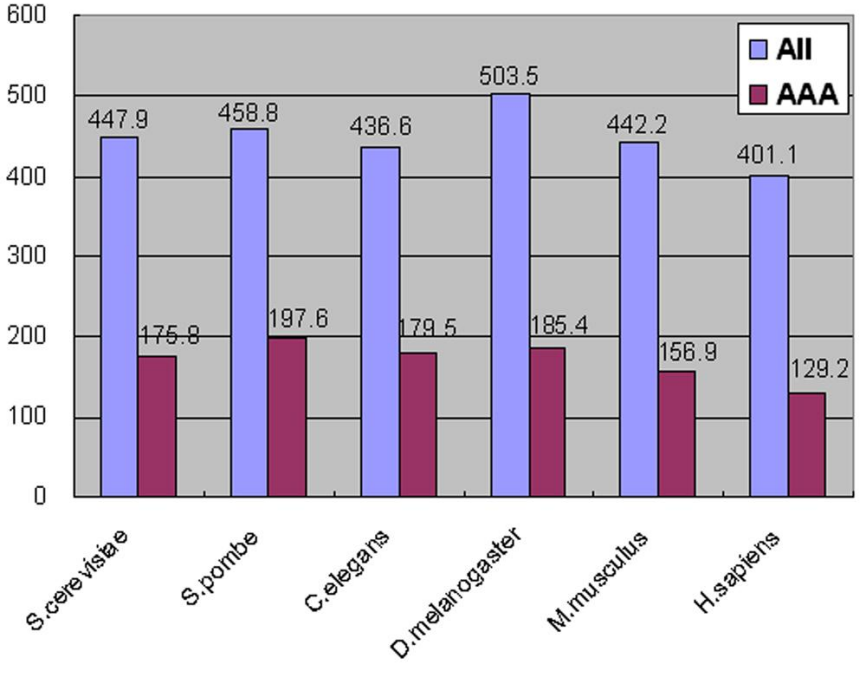


Fig. 3. The average length of AAA proteins is much shorter than the average length of all proteins. All, all proteins in an organism; AAA, AAA proteins in an organism.

Unexpectedly, AAA proteins in eukaryotes are much abundant. There are about 25.12%, 18.46%, 18.94%, 20.33%, 18.91% and 25.54% of proteomes to be AAA proteins in *S. cerevisiae*, *S. pombe*, *C. elegans*, *D. melanogaster*, *M. musculus* and *H. sapien*s. In table 2, W-type AAA proteins are the most abundant, while L- and S-type AAA proteins are much less than other types. We noticed that many proteins were fragment to be short sequence (Table 1), and removed these fragment sequences to obtain the full-length data sets. The AAA proteins in fragment and full-length proteins were also calculated. Obviously, percentiles of AAA proteins in fragment sequences are very high, from 30.45% to 62.27% (Fig. 1B). However, AAA proteins are still much ubiquitous in full length data sets, from 14.1% to 23.4% (Fig. 1A).

Theoretically, shorter proteins will be more frequent to be AAA proteins than longer proteins. Shorter proteins contain fewer amino acids, so one or several types of amino acids may be absent. Indeed, average lengths of AAA proteins in six organisms are much shorter than average length of all proteins (Fig. 3). Average lengths of all proteins in six species fluctuate from 401.1 aa to 503.5 aa, while average lengths of AAA proteins are from 129.2 aa ~ 197.6 aa.

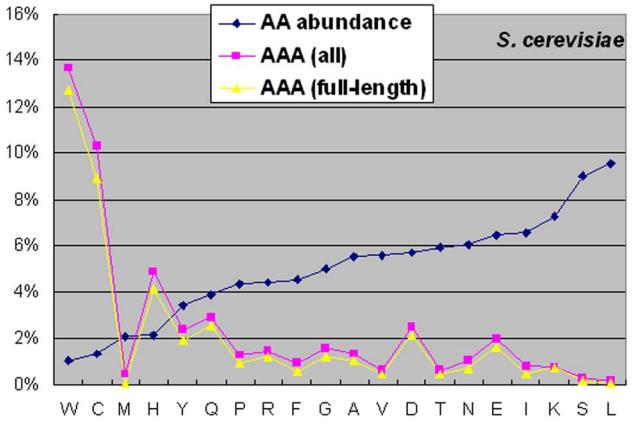
3.2 Amino Acid Absence Is Negatively Correlated with Amino Acid Abundance

To dissect the underlying mechanisms of Amino acids absence, we also calculated amino acid abundances and distributions of twenty types of residues in six organisms. In Fig. 4A and 4B, only results of *S. cerevisiae* and *H. sapiens* are shown.

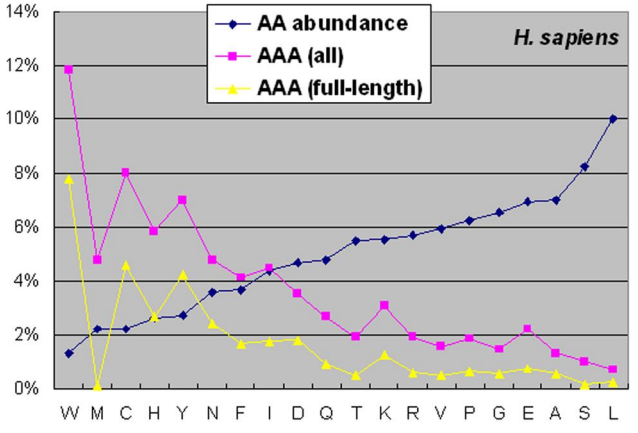
The amino acid absence is negatively correlated with amino acid abundance. And the only exceptional case is methionine (M). The M residue is usually located at the first position for transcription of proteins and coded from the standard genetic code of ATG. So at least, all of eukaryotic proteins should contain one residue at the first position. However, N-terminal M residues are often removed from mature proteins after translation. And many proteins are fragments without N-terminus. Thus, first positions of several proteins in Swiss-Prot & TrEMBL (UniProt) are not annotated as M.

In both *S. cerevisiae* and *H. sapiens*, the most abundant AAA protein is W-type AAA proteins, while its abundance and distribution is much less than other residues. The top two most abundant amino acids are S and L, while S- and L-types AAA proteins are much less than other types. The negative correlations were also found in other organisms (data not shown).

Taken together, we identified two general characteristics of AAA proteins. Eukaryotic AAA proteins tend to be very short sequences and amino acid absence is negatively correlated with amino acid abundance and distribution. Thus, we propose that most of AAA proteins are generated randomly during evolution. If a residue is not much important for a protein, it could be lost from the sequence or replaced by other similar residues, e.g., I to L and R to K, etc. Such mutations of a protein occur on its coded DNA under neutral evolution, and don't alter the function of the protein. And many short proteins lack of one or several types of amino acids might still keep their natural functions. Again, the ubiquity of AAA proteins implies that not all of twenty amino acids are indispensable to form a protein sequence.



(A)



(B)

Fig. 4. The amino acid absence is negatively correlated with the amino acid abundance and distribution. For AAA proteins, both AAA proteins in all proteins and in full-length proteins are counted. (A) *in S. cerevisiae*; (B) in *H. Sapiens.*

3.3 The Conserved AAA Proteins with Important Functions

Based on above analyses, most of AAA proteins are generated by random. However, proteins without one or several types of amino acids may also be important for their normal functions. We recalled that fifteen of the amino acids could be modified by various PTMs. And each PTM recognizes one or several kinds of specific amino acids to ensure the regulatory fidelity. Thus, a protein lack of a specific type of amino acids might prohibit some PTMs, and such mechanisms might be conserved during evolution. For example, sumoylation and ubiquitination usually modify lysine (K) residues in proteins. Thus, proteins without K residues may not be modified by sumoylation and ubiquitination. In addition, if a lysine might be less useful or even potentially

harmful for a protein's function, absence of K-amino acid will be conserved during evolution.

In this study, we considered two organisms of human and mouse to identify conserved AAA proteins. Two mammalian organisms are separated with proper evolutionary distances. So, if one or several types of amino acid absence are potentially important for proteins function, such absent events will be conserved between two organisms. With a popular software of InParanoid (ver 1.0), we identified 17,733 pairs of orthologs conserved in human and mouse. And there were 1,220 conserved AAA proteins identified (Fig. 2). Totally, there were 12,093 and 17,497 AAA proteins identified in mouse and human, respectively. Although fragment sequences were not removed in the process of orthologs detection, we can roughly estimate that ~10% of AAA proteins are conserved in human or mouse.

Table 3. There are 25 K-type AAA proteins conserved between human and mouse

Human	Mouse	Function (UniProt)
R4RL2_HUMAN	R4RL2_MOUSE	Axon regeneration
GDF1_HUMAN	GDF1_MOUSE	Cell differentiation
Q3B7T3_HUMAN	Q9EQG5_MOUSE	Unknown
BBC3_HUMAN	BBC3_MOUSE	Essential mediator of p53- dependent and p53-independent apoptosis
GP1BB_HUMAN	GP1BB_MOUSE	Formation of platelet plugs
KRA24_HUMAN	Q9D3H4_MOUSE	Interacts with hair keratins
Q5JY54_HUMAN	Q9DA47_MOUSE	Unknown
Q6UWH7_HUMAN	Q8BN06_MOUSE	Unknown
KRA42_HUMAN	Q3V4B7_MOUSE	Interacts with hair keratins
NKG7_HUMAN	NKG7_MOUSE	Integral to Membrane
KR131_HUMAN	Q8K198_MOUSE	Interacts with hair keratins
Q86SM2_HUMAN	Q8BUT3_MOUSE	Unknown
CJ035_HUMAN	CJ035_MOUSE	Unknown
Q7Z750_HUMAN	Q6PGA6_MOUSE	Unknown
MYLE_HUMAN	MYLE_MOUSE	Unknown
Q96GZ3_HUMAN	Q8BZE0_MOUSE	Mitotic cell cycle, nervous system development
KRA71_HUMAN	KRA71_MOUSE	Interacts with hair keratins
Q6IQ42_HUMAN	Q8CF51_MOUSE	Unknown
Q96B49_HUMAN	Q9CQN3_MOUSE	Unknown
KRA81_HUMAN	KRA81_MOUSE	Interacts with hair keratins
Q96AN0_HUMAN	Q8VD25_MOUSE	Unknown
Q4ZG53_HUMAN	Q6P220_MOUSE	Unknown
Q8NFP0_HUMAN	Q8K459_MOUSE	Unknown
SARCO_HUMAN	SARCO_MOUSE	Regulation of calcium ion transport
Q8IUH7_HUMAN	Q8CFD2_MOUSE	Unknown

To dissect functions of these conserved AAA proteins, we analyzed functions of twenty types of conserved AAA proteins and presented 25 conserved K-type AAA proteins in Table 3. Analyses for other types of AAA proteins are not shown due to page limitation. For 25 K-types AAA proteins, we obtained their functions from functional annotations of UniProt database, Gene Ontology annotation and literature, etc. These proteins might not be either sumoylated or ubiquinated for degradation. Totally, eleven of them are functional annotated. These AAA proteins play important roles in diverse cellular processes. For example, mouse GDF-1 (growth/differentiation factor 1) regulates left-right patterning in embryonic stem cells at early stages of development [16]. And Bcl-2-binding component 3 (BBC3), also called PUMA, is a pro-apoptotic protein and plays important roles in apoptosis of colon cancer cells [17, 18]. Its gene expression is activated by p53, then replaces p53 from Bcl-xL and leads to p53-dependent apoptosis [19], or dissociates Bax and Bax-xL to cause p53-independent apoptosis [18]. Interestingly, five Keratin-associated proteins (KRA24, KRA42, KRA131, KRA71 and KRA81) are conserved K-type AAA proteins. Keratin-associated proteins (KAPs) comprise a large superfamily, with at least 85 distinct members reported [20]. These KAPs interact with keratins to be the "glue" to hold the hair fiber together. Since the hair is hard to be degraded, it's not strange that many components of hair are hard to be degraded. In this circumstance, the K residue might be less useful for five KAPs. And conservations of five KAPs between human and mouse imply the amino absent mechanism might be also conserved.

4 Conclusions

Intuitively, proteins lack of one or several types (AAA proteins) might be much scarce, since each type of amino acid is important for a protein's function. However, our results propose that AAA proteins ubiquitously exist in eukaryotes. There are about 18.5% ~ 25.5% of proteomes in six organisms to be AAA proteins. After fragment removed, proportions of AAA protein in full-length proteins are still much high, from 14.1% ~ 23.4%. Given evidences that AAA proteins tend to be shorter sequences and amino acid absence is negatively correlated with amino acid abundance and distribution, we propose that most of AAA proteins are generated by random during evolution. However, if one type of amino acid is less useful or even harmful for a protein's function, then the amino acid absent mechanism will be conserved. Indeed, we identified 1,220 AAA proteins (~10%) to be conserved in human and mouse. These proteins play important roles in diverse functions and processes. Among 25 K-types AAA proteins, there are five Keratin-associated proteins (KRA24, KRA42, KRA131, KRA71 and KRA81) to interact with hair keratins and hold hair fibers together. The ubiquitination and degradation of these proteins are potentially not useful or even harmful for these proteins. And the K-residue absent mechanism might be conserved in these proteins.

Acknowledgements. This work was supported by grants from Chinese Academy of Science (KSCX2-YW-H 10; KSCX2-YW-R65), Chinese Natural Science Foundation (39925018, 30270654, 30270293, 90508002, 30500183, 30700138 and 60533020),

Chinese 973 Project (2002CB713700), Chinese 863 Project (2001AA215331), Chinese Minister of Education (20020358051) and the National Institutes of Health (DK56292; CA89019).

References

1. Zhou, F., Xue, Y., Yao, X., Xu, Y.: A General User Interface for Prediction Servers of Proteins Post-Translational Modification Sites. Nat. Protoc. 1, 1318–1321 (2006)
2. Seo, J., Lee, K.J.: Post-Translational Modifications and their Biological Functions: Proteomic Analysis and Systematic Approaches. J. Biochem. Mol. Biol. 37, 35–44 (2004)
3. Walsh, C.T., Garneau-Tsodikova, S., Gatto Jr., G.J.: Protein Posttranslational Modifications: the Chemistry of Proteome Diversifications. Angew. Chem. Int. Ed. Engl. 44, 7342–7372 (2005)
4. Mann, M., Jensen, O.N.: Proteomic Analysis of Post-translational Modifications. Nat. Biotechnol. 21, 255–261 (2003)
5. Olsen, J.V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., Mann, M.: Global, in Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks. Cell 127, 635–648 (2006)
6. Xue, Y., Zhou, F., Zhu, M., Ahmed, K., Chen, G., Yao, X.: GPS: A Comprehensive WWW Server for Phosphorylation Sites Prediction. Nucleic. Acids Res. 33, W184–187 (2005)
7. Blom, N., Gammeltoft, S., Brunak, S.: Sequence and Structure-based Prediction of Eukaryotic Protein Phosphorylation Sites. J. Mol. Biol. 294, 1351–1362 (1999)
8. Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., Brunak, S.: Prediction of Post-Translational Glycosylation and Phosphorylation of Proteins from the Amino Acid Sequence. Proteomics 4, 1633–1649 (2004)
9. Zhou, F., Xue, Y., Lu, H., Chen, G., Yao, X.: A Genome-Wide Analysis of Sumoylation-related Biological Processes and Functions in Human Nucleus. FEBS Lett. 579, 3369–3375 (2005)
10. Zhou, F., Xue, Y., Yao, X., Xu, Y.: CSS-Palm: Palmitoylation Site Prediction with A Clustering and Scoring Strategy (CSS). Bioinformatics 22, 894–896 (2006)
11. Li, A., Xue, Y., Jin, C., Wang, M., Yao, X.: Prediction of N(epsilon)-Acetylation on Internal Lysines Implemented in Bayesian Discriminant Method. Biochem. Biophys. Res. Commun. 350, 818–824 (2006)
12. Chen, H., Xue, Y., Huang, N., Yao, X., Sun, Z.: MeMo: A Web Tool for Prediction of Protein Methylation Modifications. Nucleic. Acids Res. 34, W249–253 (2006)
13. Catic, A., Collins, C., Church, G.M., Ploegh, H.L.: Preferred in Vivo Ubiquitination Sites. Bioinformatics 20, 3302–3307 (2004)
14. Xue, Y., Liu, D., Fu, C., Dou, Z., Zhou, Q., Yao, X.: A Novel Genome-Wide Full- Length Kinesin Prediction Analysis Reveals Additional Mammalian Kinesins. Chinese Sci. Bull. 51, 1836–1847 (2006)
15. Remm, M., Storm, C.E., Sonnhammer, E.L.: Automatic Clustering of Orthologs and Inparalogs from Pairwise Species Comparisons. J. Mol. Biol. 314, 1041–1052 (2001)
16. Rankin, C.T., Bunton, T., Lawler, A.M., Lee, S.J.: Regulation of Left-Right Patterning in Mice by Growth/Differentiation Factor-1. Nat. Genet. 24, 262–265 (2000)
17. Han, J., Flemington, C., Houghton, A.B., Gu, Z., Zambetti, G.P., Lutz, R.J., Zhu, L., Chittenden, T.: Expression of Bbc3, A Pro-Apoptotic BH3-only Gene, Is Regulated by Diverse Cell Death and Survival Signals. Proc. Natl. Acad. Sci. U.S.A. 98, 11318–11323 (2001)

18. Ming, L., Wang, P., Bank, A., Yu, J., Zhang, L.: PUMA Dissociates Bax and Bcl-X(L) to Induce Apoptosis in Colon Cancer Cells. J. Biol. Chem. 281, 16034–16042 (2006)
19. Chipuk, J.E., Bouchier-Hayes, L., Kuwana, T., Newmeyer, D.D., Green, D.R.: PUMA Couples the Nuclear and Cytoplasmic Proapoptotic Function of P53. Science 309, 1732–1735 (2005)
20. Rogers, M.A., Schweizer, J.: Human KAP Genes, Only the Half of it? Extensive Size Polymorphisms in Hair Keratin-Associated Protein Genes. J. Invest. Dermatol. 124, vii–ix (2005)

Optimal Accurate Minkowski Sum Approximation of Polyhedral Models

Xijuan Guo[1], Lei Xie[2], and Yongliang Gao[2]

[1] College of Information Science and Engineering,
Yanshan University QinHuangDao, 066004, China
xjguo@ysu.edu.cn
[2] College of Information Science and Engineering,
Yanshan University QinHuangDao, 066004, China
leilei84730@yahoo.com.cn

Abstract. This paper presents an innovational unified algorithmic framework to compute the 3D Minkowski sum of polyhedral models. Our algorithm decomposes the non-convex polyhedral objects into convex pieces, generates pairwise convex Minkowski sum, and compute their union. The framework incorporates different decomposition policies for different polyhedral models that have different attributes. We also present a more efficient and exact algorithm to compute Minkowski sum of convex polyhedra, which can handle degenerate input, and produces exact results. When incorporating the resulting Minkowski sum of polyhedra, our algorithm improves the fast swept volume methods to obtain approximate the uniting of the isosurface. We compare our framework with another available method, and the result shows that our method outperforms the former method.

Keywords: Computational geometry; Minkowski sum; Non-convex polyhera; Regular-tetrahedron Gaussian Map; Swept volume.

1 Introduction

As an important problem in the computational geometry, Minkowski sum has raised lots of attentions.Let P and Q is two closed polyhedra in R^d. The Minkowski sum of P and Q is the polyhedron $M = P \oplus Q = \{p + q \mid p \in P, q \in Q\}$, Here $p + q$ is the sum of position vectors p and q, corresponding to points contained in P and Q [1]. The problem of Minkowski sum computation has very comprehensive applications in many fields, such as computer graphics [2], CAD/CAM [3], dynamic simulation, offset operation, computing collision-free paths in robot motion planning [4].

Our goal is to compute the boundary of the Minkowski sum of two polyhedral models. The Minkowski sum of two convex polytopes P and Q, respectively with m and n features can have O (mn) combinatorial complexity, at present, many of the algorithms have been implemented which is accurate and simpler to compute, but the computation of the Minkowski sum of non-convex polyhedra have complexity as high as $O(m^3n^3)$[5] and even more worse. As a result, no practical algorithms are known for robust computation of exact Minkowski sum of complex polyhedral models.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 179–188, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2 Minkowski Sum Computation

2.1 Overview

In R^n, computing Minkowski sum of the two polyhedra is a vital geometrical operation, it can be defined as a set of pairwise sums of points from two operands respectively. It also can be described as an operand transfers along another operand's contour, in directed direction. Generally the complexityof Minkowski sum of two convex polyhedra is O (mn), which is m, n faces respectively in R^3. However, for non-convex polyhedra, the complexity is O (m^3n^3). Thus it is easier to compute Minkowski sums of convex polyhedra as compared to general polyhedral models. Recur to the following property of Minkowski sum: $S_1 \oplus S_2 = S_2 \oplus S_1$; $S_1 \oplus (S_2 \bigcup S_3) = (S_1 \oplus S_2) \bigcup (S_1 \oplus S_3)$, if $P = P_1 \cup P_2$,then $P \oplus Q = (P_1 \oplus Q) \cup (P_2 \oplus Q)$. The resulting algorithm combines this property with convex decomposition for general polyhedral models:

(1) Compute a convex decomposition for each polyhedron.
(2) Compute the pairwise convex Minkowski sums between all possible pairs of convex pieces in each polyhedron.
(3) Compute the union of pairwise Minkowski sums.

Our algorithm is based on this property too, but the perfect decomposition is not always achieved. As a result, we use different decomposing methods for different character models. The diverse convex decomposing methods have serious influence on the Minkowski sum computation. For example, we hope the number of sub convex polytopes is least, which is gained by the convex decomposition. But the least number is not most useful for improving the speed of computing Minkowski sum. The shape of polyhedron is another very important factor in this computation. In addition, although the pairwise convex Minkowski sums are convex, their complexity of the union is O (n^6) [7], n is the number of sub convex polyhedral and usually up to thousands. Computing pairwise Minkowski sums between all pairs of convex pieces results in O (n^2) pairwise Minkowski sums. We may need to compute the union of a large number of primitives (pairwise Minkowski sums), the primitives themselves are relatively simple and typically have low combinatorial complexity. Therefore, we only compute an approximate union.

2.2 Convex Decomposition

As early as 1981, Chazelle [8] et al proposed one of the earliest convex decomposition algorithms, which can generate O (r^2) convex parts and uses O (nr^3) time, where n and r the number of polygons and notches in the original polyhedron. However, no robust implementation of this algorithm is known. Most practical algorithms for convex decomposition are surface decomposition or tetrahedral volumetric decomposition in Refs. [8,10]. Typically, these methods can generate O (n) convex parts and each of them has a few faces.

There are two type polyhedra in our real word, curved objects and solid objects. We should employ different methods to decompose them. In this way, many computational operations are conveniently carried out, in this text, the operation is referred

to compute Minkowski sum of polyhedral models. Now we describe the tetrahedral decomposition and surface decomposition respectively.

2.2.1 Tetrahedral Decomposition

Let P is a solid polytope in R^3, P is decomposed into non-intersecting tetrahedrons, all points of tetrahedrons is belonged to primitive polytope P, this decomposition is called tetrahedralization. This decomposition is very challenge compared to corresponding planar problems. In 1981,Chazelle point out : for any suffice big n, there are a simple polytope containing n vertical points, and $\Theta(n^2)$ additional vertical points, which is called steiner points, as long as these steiner points are admitted in the decomposition, the tetrahedralization must be implemented for any simple polytope. Later, Chazelle and Palios [12] in 1990 improve this ambit to $\Theta(n+r^2)$, r is the reflex edge of polytope, and the run time of this decomposition is $O(nr+r^2 \log r)$.

We used a modification of the convex decomposition scheme. We show the main step to implement the tetrahedralization. The procedure of the algorithm is described as follows:

(1) Input the coordinates values of solid polyhedron's vertex, the list of points coordinates composing the plane, that is the data structure of the polyhedron, S;
(2) According to planar equation, we can get the function of the Polygon (*A, B, C)*, and get all of planar equations composing the polyhedron;
(3) Choosing a convex point from the list of points, all its abutting vertexes are projected to the plane, then use Delaunay triangulation to get the convex space of the convex point;
(4) This convex point and its convex space are separated from the primitive polyhedron, becoming a sole tetrahedron;
(5) Modify the data structure of polyhedron S;
(6) Judge whether the decomposition is finished, if not, return (4); if finished, return true.

2.2.2 Surface Decomposition

The perfect decomposition is not implemented exactly for curved polyhedra [10]. In that case, other techniques should be applied. We use the surface decomposition presented by Ehmann, this methods can generate O (n) convex parts and each of them has a few faces. This method is available in a public collision detection library, SWIFT++ [13].Let S is an orientable closed polyhedral surface S, the surface of polyhedral models P is $S=\partial P$, the surface decomposition can be defined as:

$$(S \bigcap C_i) = c_i , \quad \partial P = \bigcup c_i, \quad \forall\, i,j: i \neq j,\ c_i \cap c_j = \varnothing ,\ c_i \subseteq \partial (C_i) \qquad (1)$$

Where c_i are convex surface patches that lie entirely on the surface (boundary) of their convex hull, which is the corresponding convex piece. The decomposition for P is composed of the convex surface C_i, the pieces are contained entirely within P but their union does not necessarily equal P. The convex surface C_i consists of two types of faces: real faces that belong to the original polyhedron and virtual faces that are

artificial of the convex hull computation. We disregard these virtual faces though building up a record for these faces when we compute them Minkowski sum.

2.3 Pairwise Minkowski Sum Computation

We use Regular Tetrahedron Gaussian Map algorithm to compute convex polyhedra Minkowski sum. We compute the pairwise Minkowski sums between all possible pairs of convex pieces, P_i and Q_j belonging to P and Q, respectively. Let us denote the resulting Minkowski sum as M_{ij}. Our algorithm is based on the Gaussian Map, we now introduce the Gaussian Map compactly, and then describe the RTGM algorithm in detail.

2.3.1 The Regular Tetrahedron Gaussian Map

Similar to the Gaussian map [14], the Regular Tetrahedron Gaussian Map (RTGM) T of a convex polytope P in R^3 is a set-valued function from P to the four faces of the unit tetrahedron. A facet f of P is mapped under T to a central projection of the outward unit normal to f onto one of the tetrahedron faces. Observe that, a single edge e of P is mapped to one or a chain of connected segments that lie in a or three abutting tetrahedron-faces respectively, and a vertex v of P is mapped to at most three abutting convex dual faces that lie in three abutting cube-faces respectively. According to the above rule, we can get RTGM of convex polyhedron. Then, we simplify the construction and manipulation of the representation, as the subdivisions of each tetrahedron face is a planar map of segments. We use the CGAL Arrangement-2 data structure to maintain the planar maps. The construction of the four planar maps from the polytope features and their incidence relations amounts to the computation of the overlay of four pairs of planar maps, an operation well supported by the data structure as well.

Each facet of a RTGM is a convex subdivision, and the inverse RTGM, denoted by T^{-1}, maps the four faces of the unit tetrahedron to the polytope boundary. Each planar face f is extended with the coordinates of its dual vertex $v = T^{-1}$ (f) among the other attributes (detailed below), resulting in a unique representation.

Doubly Connected Edge List (DCEL) data structure [15] is a representation based on edges, and includes the information of vertices and faces, each topological edge of the subdivision is represented by two halfedges with opposite orientation, and each halfedge is associated with the face to its left. DCEL contains a record for each face, halfedge and vertex of the subdivision. The record that we define for every vertex is: the coordinate of this vertex, one of the incident edges whose origin is this vertex; the record for every halfedge is: the origin of this halfedge, the twin of this halfedge, the next of this halfedge; the record for every face is: one of halfedges on its outer boundary, so that we can gain the incident relation among faces, halfedges and vertices in each planar subdivision. There is some additional information in the planar subdivision features, which is very important for the Minkowski sum computation, we define the attribute information for every vertex and every halfedge.

The exact mapping from a facet normal in the R^3 coordinate system to a pair that consists of a planar map and a planar point in the R^2 coordinate-system is defined precisely through the correctly choose the coordinate system, illustrated in Fig. 1(a). Here we define that the id of the corner vertex which lies in the positive orientation of axes y' is 0, and the id of the corner vertex the lies in the positive orientation of axes

Table 1. Cyclic Traverse Table of Corner Vertices

The relation between f and coordinate		Corner ID 0		1		2	
ID	3D	Face ID	CV ID	Face ID	CV ID	Face ID	CV ID
Ⅰ	z<0	Ⅱ	1	Ⅲ	1	Ⅳ	1
Ⅱ	x>0	Ⅳ	0	Ⅲ	2	Ⅰ	2
Ⅲ	x<0	Ⅱ	0	Ⅳ	2	Ⅰ	0
Ⅳ	y<0	Ⅲ	0	Ⅱ	2	Ⅰ	1

CV stands for Corner Vertices.

x' is 1, illustrated in Fig. 1(b). We have a fast clockwise traversal of the faces incident to any given vertex v, if v is an internal vertex, clockwise traversal around v is immediately available by the DCEL, if v is a boundary vertex, clockwise traversal around v is enabled by the cyclic table in Table 1.

After introduced the Regular Tetrahedron Gaussian Map and some assistant operations for computing Minkowski sum, we would describe the algorithm of computing exact Minkowski sum computation.

2.3.2 Exact Minkowski Sum Computation

The overlay of two planar subdivisions φ_1 and $\varphi_2$is a planar subdivision φ, such that there is a face f in φ if and only if there are faces f1 and f2 in φ_1 and φ_2 respectively such that f is a maximal connected subset of $f_1 \cap f_2$. The whole algorithm includes three parts:

(1) Compute overlay of four pairs of planar subdivisions, containing three steps: computing intersection, reconstructing topology and constructing the DCEL for overlay graph;

(2) Fulfill the distribution for attributes, that is updating the record of the DECL;

(3) Get the Minkowski sum polyhedron though computing T^{-1}.

At present, the most well-know algorithm for computing polyhedral Minkowski sum is the convex hull[11] and the cubical Gaussian Map that produce exact results. The convex hull is based on the definition of the Minkowski sum: $P \oplus Q=CH(\{v_i+v_j | v_i \in V_p, v_j \in V_q\})$, here CH denotes the convex hull operator, V_p, Vq represent the sets of vertices of the polyhedron P and Q. Let both of the two polyhedra have n vertices, then their sets of vertices is $O(n^4)$ at worst case in R^3 when computing their Minkowski sum using convex hull approach. The Cubical Gaussian Map CGM algorithms use a dual representation of convex polyhedra based on Gaussian Map. A convex polytope P in R^3 is mapped to the unit cube whose edges are parallel to the major axes and are of length two under the CGM, and the six faces of the cube is seen as six planar subdivisions. As a result, the construction of the Minkowski sum amounts to the computation of the overlay of six pairs of planar maps, an operation well supported by the DECL data structure as well.

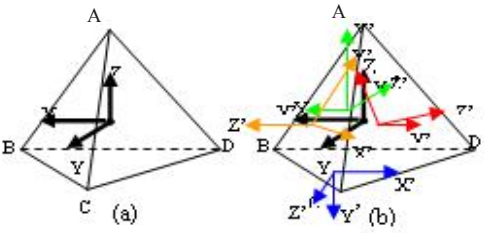


Fig. 1. (a) Building the main coordinates system for tetrahedron. (b)building the coordinate system for each face.

So our algorithm not only can get exact convex polyhedra Minkowski sum, but also simper implement. The experiment results indicate that our algorithm is faster than the convex hull and CGM approach .Table2 shows the results produced by experiments and compares the three algorithms.

Table 2. Time consumption of the Minkowski-sum computation (seconds)

Object type		Minkowski sum								
		Prima			Dual			CGM	CH	RTGM
Obj_1	Obj_2	V	E	F	V	HE	F			
Octa	Octa	6	12	8	24	48	12	0.008	0.012	0.005
Icos	Icos	12	30	20	72	192	36	0.030	0.130	0.008
PH	Ell16	542	1102	643	978	2922	560	1.117	14.024	0.808
Mush	RGS4	1254	2516	1504	1378	10256	1642	3.126	136.692	1.642

Octa-Octahedron;Icos-Icosahedron;PH-PentagonalHexecontahedron; Ell16-Ellipsoid-like polyhedron made of 16 latitudes and 32 longitudes; Mush-Polyhedron-like mushroom; RGS4-Geodesic Sphere level 4.

3 Union Computation and Minkowski Sum Approximation

The accuracy of the algorithm is mainly governed by the resolution of the underlying grid and the reconstruction algorithm. If the final surface has thin features, insufficient grid resolution can create handles in the reconstructed surface. Moreover, many geometric operations create new sharp features or edges on the boundary of the final surface. The accurate union (M is the union of pairwise sub polyhedra: $M=\bigcup M_{ij}$), is too waste time, and the existent algorithms are either not robust or too slow in terms of computing the union of a high number of polyhedral models [17]at present As a result, instead of computing M exactly, we compute M' using distance field-based techniques. This overall procedure is composed of the following steps:

(1) Sampling: Generate an externally tangent cubic grid for sub Minkowski sum polyhedra respectively and compute the signed distance field at its grid points.
(2) Operation: For each geometric operation (e.g., union), perform a min/max operation on the signed distance field of the primitives.

(3) Reconstruction: Use Enhanced Marching Cubes (EMC) [17] to perform isosurface extraction from the distance field.
(4) Topological refinement: In view of Swept Volume (SV), the extracted isosurface should be checked, in order to ensure a single connected isosurface. The extracted isosurface is our approximation to the boundary of the Minkowski sum.

Trilinear interpolation is a common way to define a function inside each cell of the grid. It is computationally simple and provides good estimation of a function between sampled points in practice. Isosurface extraction [16] is one of the most common techniques for visualizing the volumetric data. An isosurface is a level set surface defined as:

$$I(w) = \{(x, y, z) \mid F(x, y, z) = w\} \tag{2}$$

where F is a function defined from the data and *w* is an isovalue. The isosurface I is polygonized for modeling and approximation boundary.

Our method is to set up externally tangent cubic grid for sub Minkowski sum polyhedra. In this way, many test works have been removed and this criterion ensures that the surface intersects the grid cell in a simple manner in most cases. Early Marching Cubes algorithm performs isosurface extraction from the distance field producing artifacts aliasing[9], and requiring noting the relations between all vertexes so that algorithm needs big memory. Be against these disadvantages, some variant excellent algorithm have been presented [17]. Thus we have a better approximation to the exact, outer boundary. Our algorithm generates a closed, connected polygon structure. Finally, the topological guarantees [18] ensure a good quality of the Minkowski sum approximation.

4 Implementation and Performance

In this section, we describe the implementation of our approximation algorithm, and demonstrate its performance on different polyhedral models, in the end, we compared our algorithm with the former algorithm.

The experiment condition: CPU: Pentiu2.8GHz; Main memory, 1G; Operating system: Windows XP; Software Library: LEDA (Library of Efficient Data Types and Algorithms), providing exact real number [19]. CGAL (Computer Geometry Algorithm Library). PVL (Polyhedral Mode Vertex Library) denoting the polyhedral models, recording the verticals coordinates value of the polyhedra. Our algorithm is simple to implement and the topological guarantees ensure a good quality of the Minkowski sum approximation.We tested our algorithm on a number of complex models. The whole frame of our algorithm for computing Minkowski sum of two polyhedra is presented below.

Algorithm Compute Minkowski sum (*P, Q*)

Begin

Step1. Identify the type of the P and Q; if both of them is convex solid polyhedra, please turn to the Step5, else do the Step2;

Step2. If at least one of them is non-convex solid polyhedra, and then turns it to the Step 3, else if at least one of them is non-convex curved polyhedra, and then turns it to the Step4;

Step3. Tetrahedral Decomposition for non-convex solid polyhedra, then to the Step5;

Step4. Surface Decomposition for non-convex curve polyhedra, then to the Step5 too;

Step5. Compute the Minkowski sum of the two convex polyhedra by the RTGM;

Step6. Unite the pairwise Minkowski sum of resulting M_{ij} using fast SV method.

End.

The pivotal issue is the decomposition in the above process, Tetrahedral Decomposition and Surface Decomposition. The RTGM is a novel algorithm for computing Minkowski sum of two convex polyhedra. In the end, the fast SV method is improved using establishing externally tangent cubic grid.Table3 shows the performance of our algorithm and the compare with the former method presented by G.Varadhan, D. Manocha in 2005[11].

Table 3. The performance of our algorithm on different models and the compare with G.V

Model1		Model 2		the total	Performance（seconds）		
name	convex number	name	convex number	number of con-number	Convex	Union Minkowski	G.V
Spoon	78	Sphere	1	1	0.58	21.348	26.004
Light	24	Spring	133	165	0.76	27.172	29.842
Cup	338	Sphere	1	338	1.04	28.036	32.200
Knife	273	Scissors	230	62,790	34. 82	736.64	778.06

5 Path Planning Application

We describe an application of our approximate algorithm. Certainly, our alogrithm also can be applied in morphological operations, and penetration depth computation, and so on.Path planning is an important problem in algorithmic robotics. The basic problem is to find a collision-free path for a robot among rigid objects [14]. We calculate the case of a 3D polyhedral robot undergoing translation motion among 3D polyhedral obstacles. This problem is often formulated using a configuration space approach. The free configuration space is the set of all possible positions in which the robot avoids contact with the obstacles. It can be expressed as the complement of the Minkowski sum of the robot and the obstacles.In order to find some viable pathes, it ables to have a representation that captures the connectivity of the free configuration space. Fig. 2 shows application of our algorithm to assembly planning. It consists of two parts each with a screw cap and nut. The goal is to assemble the two parts so that the nut fit into the screw cap.

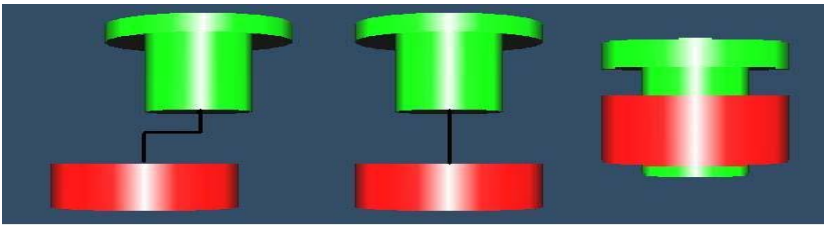


Fig. 2. Assembly planning: it shows application of our algorithm to assembly planning. The three images on the left shows a path that the robot can take so that the two parts could be assembled. The rightmost image shows the Minkowski sum and the path of the robot in configuration space.

6 Conclusion and Future work

We have improved the unified algorithmic framework to approximate the 3D Minkowski sum of polyhedral objects. Exploiting externally tangent techniques, the performance of fast swept volume algorithm is improved. Our algorithm guarantees that the approximation has the correct topology on the approximation. As part of future work, we would like to apply our algorithm to more practical applications such as penetration depth computation, exact collision detection. We would like to develop better convex decomposition algorithms which can produce smaller numbers of convex objects. It is well-known that the Minkowski sum of two star-shaped polyhedral is a star-shaped polyhedron. We could exploit this property and design our overall approach based on star-shaped decomposition instead of convex decomposition. The main advantage of that is that the star-shaped decomposition of a polyhedron would typically result in fewer primitives.

References

1. Ghosh, P.: A Unified Computational Framework for Minkowski Operations. Comp. Graph 17(4), 357–378 (1993)
2. Badlyans, D.: A Computational Approach to Fisher Information Geometry with Applications to Image Analysis. In: Proceedings of 5th International Workshop, pp. 18–33 (2005)
3. Yun, S.C., Byoung, C., Choi, K.: Contour-parallel Offset Machining Without Tool-Retractions. J. Computer-Aided Design 35(9), 841–849 (2003)
4. Lozano, T.: Spatial Planning: A Configuration Space Approach. IEEE Trans. Comput. C-32, 108–120 (1983)
5. Chazelle, B.: Convex Partitions of Polyhedra: A Lower Bound and Worst-case Optimal Algorithm. SIAM J.Comput. 13(7), 488–507 (1984)
6. Fogel, E., Halperin, D.: On the Exact Maximum Complexity of Minkowski Sums of Convex Polyhedra. In: Proceedings of 23rd Annual ACM Symposium on Computational Geometry, pp. 319–326 (2007)
7. Chazelle: Convex Decompositions of Polyhedra. In: ACM Symposium on Theory of Computing, pp. 70–79 (1981)
8. Scott, S., Tao, J., Joe, W.: Manifold Dual Contouring. J. IEEE Transctions on Visualization and Computer Graphics 13(3), 610–620 (2007)

9. Chazelle, B., Dobkin, D., Shouraboura, N.: Strategies for Polyhedral Surface Decomposition: An Experimental Study. J. Comput. Geom. Theory Appl., 327–342 (1997)
10. Varadhan, G., Manocha, D.: Accurate Minkowski Sum Approximation of Polyhedral Models. In: Proc. Comput. Graph and Appl, 12th Paci_c Conf. on (PG 2004), pp. 392–401. IEEE Comput.Sci., Los Alamitos (2005)
11. Chazelle, B., Palios, L.: Triangulating a Non-convex Polytope. Discrete Comput. Geom., 505–526 (1990)
12. Ehmann, S., Lin, M.C.: Accurate and Fast Proximity Queries Between Polyhedra Using Convex Surface Decomposition, Comput. Graph. Forum. In: Proc. of Eurographics 2000, vol. 20(3), pp. 500–510 (2001)
13. Fogel, E., Halperin, D.: Exact and Efficient Construction of Minkowski Sums of Convex Polyhedra with Applications. Computer-Aided Design 39, 929–940 (2007)
14. Computer Geometry Algorithm Library, http://www.cgal.org
15. Kim, K., Nakhoon,: Fast Extraction of Polyhedral Model Silhouettes From Moving Viewpoint on Curved Trajectory. Computers & Graphics 29, 393–402 (2005)
16. Chiou, J., Clay, V.: An Enhanced Marching Cubes Algorithm for In-Process Geometry Modeling. Manufacturing Science and Engineering 129(3), 566–574 (2007)
17. Sohn, B.: Topology Preserving Tetrahedral Decomposition of Trilinear Cell. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4487, pp. 350–357. Springer, Heidelberg (2007)
18. Library of Efficient Data Types and Algorithms, http://www.mpi-inf.mpg.de

Fuzzy C-Means Based DNA Motif Discovery

Mustafa Karabulut and Turgay Ibrikci

Department of Electrical and Electronics Engineering
Çukurova University
Adana, Turkey
mkarabulut@gmail.com, ibrikci@cukurova.edu.tr

Abstract. In this paper, we examined the problem of identifying motifs in DNA sequences. Transcription-binding sites, which are functionally significant subsequences, are considered as motifs. In order to reveal such DNA motifs, our method makes use of Fuzzy clustering of Position Weight Matrix. The Fuzzy C-Means (FCM) algorithm clearly predicted known motifs that existed in intergenic regions of GAL4, CBF1 and GCN4 DNA sequences. This paper also provides a comparison of FCM with some clustering methods such as Self-Organizing Map and K-Means. The results of the FCM algorithm is compared to the results of popular motif discovery tool Multiple Expectation Maximization for Motif Elicitation (MEME) as well. We conclude that soft-clustering-based machine learning methods such as FCM are useful to finding patterns in biological sequences.

Keywords: Motif Finding, Fuzzy Clustering, DNA Sequences.

1 Introduction

DNA is the key molecule in a cell that has responsibilities in important cell activities such as protein synthesis. Genes which make up a DNA sequence are used as templates for proteins in the "transcription" process. Understanding regulatory elements of the transcription process is one of the most challenging subjects in Bioinformatics. A transcription binding site is such a functionally important short DNA sequence, which plays a regulatory role in the transcription. These functionally important and overrepresented sequence patterns are generally 5-20 base pair (bp) long. The finding of these motif patterns is not a straightforward task since the patterns are fairly short, may have mutations, insertions and deletions [1].

There are many algorithms to detect overrepresented motifs using probabilistic, word enumerative and dictionary based methods. Well-known probabilistic methods are Multiple EM for Motif Elicitation (MEME) [2] which uses Expectation Maximization (EM) and Gibbs Sampling, a Monte-Carlo procedure. Recently, unsupervised machine learning algorithms are catching on. One example of such algorithms is using Self-Organizing Map (SOM) for the clustering phase of the motif-discovery algorithm. Mahony et al. [3] studied Self-Organizing Map of Position Weight Matrices (PWMs) to discover motifs in DNA sequences. In another study [4], SOM was used with hierarchical structure to predict motifs. These methods, which mainly rely on

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 189–195, 2008.
© Springer-Verlag Berlin Heidelberg 2008

SOM, are reported to be successful and advantageous in some ways with respect to existing algorithms.

In this paper, another machine learning technique, Fuzzy Clustering, is employed to detect overrepresented motifs in regulatory regions. Fuzzy C-Means (FCM) is a proven method in pattern recognition. This study makes the use of FCM algorithm to cluster Position Weight Matrices (PWMs) of DNA subsequences as motif candidates. PWM, which consists of frequency of a nucleotide in a specific position, is a popular method to represent motifs [5].

2 Dataset and Methods

2.1 Datasets

The datasets of this study is obtained from "The Promoter Database of Saccharomyces cerevisiae" web site [6] and "The Fraenkel Lab" web site [7]. The datasets contain the promoter regions of GAL4, CBF1 and GCN4 transcription regulators on which transcription binding sites are known to reside. To find motifs through the sequences, a background model is needed. Thus, we used the whole dataset of intergenic regions of the yeast genome to produce the background model.

These datasets are frequently used for evaluating motif discovery algorithms by many researchers.

2.2 Method

The method predicts the motifs by clustering subsequences of a given DNA dataset. When motifs are searched, each subsequence is taken at length *m*, as *sliding windows*. In training phase, each window (actually a motif candidate of length *m)* is handled by the FCM clustering algorithm and assigned to a relevant cluster according to maximum log-likelihood score of the string. Having completed the training phase, the motif candidates at each cluster center are ranked by considering "Z-Score"s. A "Z-Score" of a cluster center means that whether the motif candidate, being represented by the PWM at the cluster, is statistically interesting or not. The cluster center with the highest Z-Score is the predicted motif candidate of the biologically known motif pattern.

2.2.1 Position Weight Matrix

Position Weight Matrix (PWM), also known as Position Specific Scoring Matrix (PSSM), is one of the most popular means of motif representation for finding DNA motifs. A PWM is the core of each cluster that holds position specific information content generated from the subsequences that are members of the cluster. Information content of a specific nucleotide at a specific position in the string is calculated with Equation 1:

$$W_{ib} = \log\left(\frac{f_{ib}}{p_b}\right) \tag{1}$$

where f_{ib} is the frequency of the letter b (Nucleotides: A, C, G or T) at position i and p_b is the background frequency of the letter b. Thus, a similarity score which will be used to compare a subsequence window to a cluster center should be as follows:

$$S(x) = \sum_{i=1}^{l} \sum_{b=a}^{T} x_{ib} \log\left(\frac{f_{ib}}{p_b}\right) \tag{2}$$

The Equation 2 gives us the score of how much a given string resembles a PWM matrix. These scores will be used to calculate the membership matrix which is a requirement of the FCM algorithm for a given sequence window.

2.2.2 Structure of Fuzzy C-Means for PWM Clustering

Fuzzy C-means algorithm is an unsupervised clustering method. It simply relies on gradually moving cluster centers closer to values in the input space until a satisfactory clustering of each input is accomplished. In contrast to hard-clustering methods, each input vector can be a member of more than one cluster. A membership matrix of x by y, where x is the number of clusters and y is the number of input vectors, is used to specify how much a vector from input space is member of each cluster center. The membership matrix that should be updated in each iteration of the training can be obtained as:

$$u_{ij} = \frac{\left(\frac{1}{d^2(X_j - V_i)}\right)^{1/(q-1)}}{\sum_k \left(\frac{1}{d^2(X_j - V_i)}\right)^{1/(q-1)}} \tag{3}$$

where X denotes input vectors and V denotes cluster centers. Performance of the FCM algorithm is dependent on the q parameter, which specifies the *fuzziness* quantity. In each training iteration, the cluster centers move towards the centers of the input vectors that can be shown as follows:

$$V_i = \frac{\sum_J (u^{ij})^q X_j}{\sum_J (u^{ij})^q} \tag{4}$$

The original FCM algorithm is slightly modified for the motif discovery problem. In general applications of the FCM algorithm, the distance is usually Euclidean, whereas in our application it is the result of the scoring function, $S(x)$, given in Equation 2. Each input vector, for the motif problem that is a string of DNA nucleotides, is taken sequentially from the collection of sliding DNA sequence windows. These strings are compared against the cluster centers which are nodes on a wxh grid, where w and h specify the width and the height of the grid. The number of cluster

centers is *a priori* for FCM algorithm. *Wxh* is the number of cluster centers, which should be given in proportion to the length of DNA sequence in the study. This relation between grid size and the length of dataset is studied by Mahony et al. [3] and given quantitatively in their article.

2.2.3 Z-Score

In order to eliminate statistically uninteresting motif candidates, a frequently used method is Z-Score, or *Standard Score*. A statistically uninteresting motif candidate is a subsequence of DNA sequence which could be coincided by chance. In other words, it is a subsequence similar to the background model. The background model is obtained by generating Hidden Markov Model (HMM) of the genome dataset which our DNA sequences belong to. Generally using a 3rd or 4th order HMM is used as a background model in motif discovery applications. Z-Score is provided by Equation 5 [8]:

$$Z - Score_j = \frac{Oj - Ej}{\sigma j} \tag{5}$$

where j is the cluster, Oj is the number of sequence occurrences at the cluster and Ej is the expected number of occurrences at the cluster. Artificial DNA sequences are generated by taking into account the background model to calculate the standard-deviation and the expected number of occurrences. The artificial DNA sequences are given to the clustering system the same way with the real DNA dataset. Briefly, the more artificial DNA subsequences coincide with a node, the less is the Z-score of the node.

3 Discussion

3.1 Implementation Results

Since the number of cluster centers is *a priori* for the FCM algorithm, firstly the number of cluster centers in a grid structure should be decided. As described in Section 2.2.2, the number of cluster centers is directly proportional to the length of the dataset that is the number of nucleotides in the sequences. Thus, different grid sizes are preferred for each dataset as shown in Table 1. The algorithm is trained with different number of epochs to observe the effects of the training times on predicted motif quality. The number of epochs showed that up to 100 epochs, more training is better, but training more than 100 epochs does not affect the motif quality very much.

With these parameters applied, the algorithm clearly predicted the known GAL4, CBF1 and GCN4 motifs as can be seen in sequence logos provided in Fig. 1 and Fig. 2. The sequence logos in Fig. 1 are produced by using the known sites in the dataset.

Table 1. Selected grid size for each dataset used

Dataset	Length(bp)	Grid Size
GAL4	3246	20x10
GCN4	9488	40x30
CBF1	4690	20x20

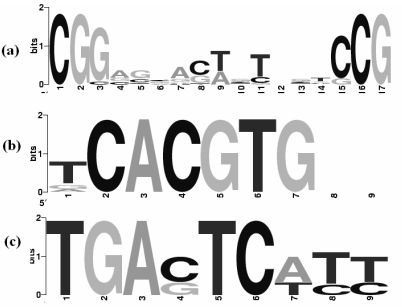


Fig. 1. The known GAL4 (a), CBF1 (b) and GCN4 (c) motifs

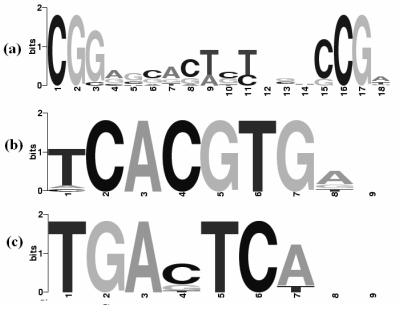


Fig. 2. The predicted GAL4 (a), CBF1 (b) and GCN4 (c) motifs by FCM algorithm

3.2 Comparison

We also implemented the method with other clustering algorithms, such as SOM and K-Means, with appropriate parameters and the same datasets given. The SOM algorithm performed well and predicted the most significant parts of the known motifs (Fig. 3).

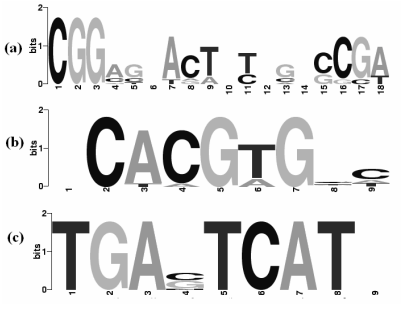


Fig. 3. The results of the SOM based implementation

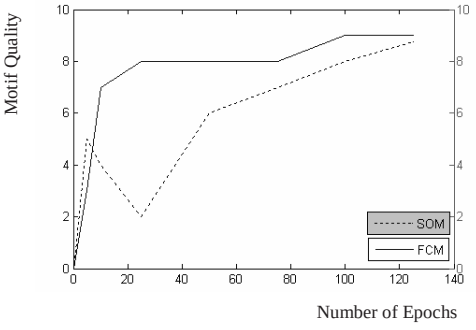


Fig. 4. Comparison of SOM and FCM in terms of motif prediction success

If FCM and SOM are compared in terms of predicted motif quality the FCM algorithm we've implemented performed slightly better than the SOM based implementation. On the other hand, as can be seen in Fig. 4, FCM algorithm converges to a qualified motif prediction at lower epochs than SOM does. The motif quality is directly proportional with clustering accomplishment of the algorithm. If all the known sites of a given motif are clustered into the same cluster center, the algorithm will give a pattern prediction very similar to the original motif. Fig. 4 explains that FCM algorithm needs less training than SOM does to gather the known sites into the same cluster.

As compared to SOM and FCM, K-means performed worse than both of FCM and SOM. As a result, both FCM and SOM as soft-clustering methods performed well and a being hard-clustering method, K-means did not produce a reasonable prediction. Thus, it may be said that soft-clustering methods are more appropriate for biological pattern search problems such as motif discovery.

Also, we compared the results of the FCM algorithm with MEME that is the well-known motif discovery software [9]. The results of MEME for the given dataset can be seen in Fig. 5. The sequence logo in Fig. 5 shows us the FCM algorithm produced more precise results than MEME did. This study showed that FCM algorithm is capable of successfully dealing with DNA sequences to reveal overrepresented motifs.

Fig. 5. The results of online MEME software with our datasets given

4 Concluding Remarks

The FCM algorithm predicted known motifs in GAL4, CBF1 and GCN4 intergenic DNA sequences of *Saccharomyces cerevisiae*. With appropriate background provided, the method is expected to predict all DNA sequence based motifs of various lengths. Also, this study shows that soft-clustering based algorithms which assign a data input to multiple clusters with a criteria such as *membership* or *neighborhood*, are suitable for clustering purposes of biological sequences.

As a future work, we will try to improve our algorithm by using recently popularity gaining methods such as employment of *Phylogenetic information*. Phylogenetic information may help to select highly conserved motifs [10], thus a combination with FCM algorithm gives hope for better results.

References

1. Das, M.K., Dai, H.K.: A Survey of DNA Motif Finding Algorithms. BMC Bioinformatics 8 Suppl 7, 21 (2007)
2. Timothy, L.B., Nadya, W., Chris, M., Li, W.W.: MEME: Discovering and Analyzing DNA and Protein Sequence Motifs. Nucleic Acids Research 34, 369–373 (2006)
3. Mahony, S., Hendrix, D., Smith, T.J., Golden, A.: Self-Organizing Maps of Position Weight Matrices for Motif Discovery in Biological Sequences. Artificial Intelligence Review 24, 397–413 (2005)
4. Derong, L., Xiaoxu, X., Bhaskar, D., Huaguang, Z.: Motif Discoveries in Unaligned Molecular Sequences Using Self-Organizing Neural Networks. IEEE Trans. Neural Networks 17(4), 919–928 (2006)
5. Sandve, G.K., Drablos, F.: A Survey of Motif Discovery Methods in an Integrated Framework. Biol. Direct. 1, 11 (2006)
6. The Promoter Database of Saccharomyces cerevisiae (SCPD) web site, `http://rulai.cshl.edu/SCPD/`
7. The Fraenkel Lab web site, `http://fraenkel.mit.edu/`
8. Ferreira, P.G., Azevedo, P.J.: Evaluating Protein Motif Significance Measures: A Case Study on Prosite Patterns. In: Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining CIDM (2007)
9. MEME (Multiple EM for Motif Elicitation) web site, `http://meme.sdsc.edu/meme/meme.html`
10. Shane, T., Jensen, S.L., Liu, J.S.: Combining Phylogenetic Motif Discovery and Motif Clustering to Predict Co-regulated Genes. Bioinformatics 21(20), 3832–3839 (2005)

Find Key m/z Values in Predication of Mass Spectrometry Cancer Data

Yihui Liu[1] and Li Bai[2]

[1] School of Computer Science and Information Technology,
Shandong Institute of Light Industry, Jinan, Shandong, China, 250353
yxl@sdili.edu.cn
[2] School of Computer Science, University of Nottingham,
Jubilee Campus, Nottingham, UK, NG8 1BB

Abstract. To find the significant biomarker is very important in detecting protein patterns associated with diseases. In this study multilevel wavelet analysis is performed on high dimensional mass spectrometry data to extract the detail coefficients, which are used to detect the difference between cancer tissue and normal tissue. In order to find the key m/z values of mass spectra, wavelet detail information is reconstructed based on orthogonal wavelet detail coefficients, and genetic algorithm is further employed to select best features from the reconstructed detail information. Finally the corresponding significant m/z values of mass spectra are identified using the optimized detail features.

1 Introduction

To find the significant biomarker is very important in detecting protein patterns associated with diseases. Surface enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF-MS) have widely been used for the detection of early-stage cancer [1,2,3]. Advanced data mining algorithms are employed to detect protein patterns associated with diseases.

Lilien et al. [4] extract features of mass spectra using principle component analysis (PCA) and linear discriminant analysis (LDA), and construct nearest centroid classifier. In the study of Wu et al.[5], T-statistics rank key features of data, support vector machines (SVM), random forests, linear/quadratic discriminant analysis (LDA/QDA), knearest neighbors, and bagged/boosted decision trees are employed to classify the data. Both the genetic algorithm (GA) approach and the nearest shrunken centroid approach have been found inferior to the boosting based feature selection approach in [6]. Levner [7] investigates the performance of the nearest centroid classifier using the following feature selection algorithm; Student-t test, Kolmogorov-Smirnov test, and the P-test are univariate statistics used for filter-based feature ranking; Sequential forward selection and a modified version of sequential backward selection are also tested. Embedded approaches include shrunken nearest centroid and an improved boosting based feature selection. Dimensionality reduction approaches, such as principal component analysis and linear discriminant analysis are tested. Yu et al. [8] develop a hybrid method for dimensionality reduction and test on a published ovarian high-resolution SELDI-TOF dataset. They use a four-step strategy for data preprocessing. Approximation coefficients

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 196–203, 2008.
© Springer-Verlag Berlin Heidelberg 2008

at first level based on wavelet analysis are used to characterize features of mass spectra. In our previous research [9], we perform multilevel wavelet analysis on mass spectra, detail coefficients measure the difference between cancer tissue and normal tissue, and the experiments prove that detail coefficients at 3rd, 4th, and 5th level achieve good results. The detail coefficients at 3rd level obtain best performance, removing the "small change" and not losing too much information. Genetic algorithm is performed to select the best features from both detail coefficients and approximation coefficients at 3rd level [10]. Experimental results suggest that both approximation coefficients and detail coefficients contribute to the classification performance of mass spectra.

In this study we extend our previous research. In order to find significant key m/z values of mass spectra, we reconstruct wavelet details from orthogonal wavelet detail coefficients at 3rd level, and reconstructed details reflect the information of original data space. Then we perform genetic algorithm on reconstructed detail to select the best features. Finally the corresponding m/z values are identified based on the optimized wavelet detail features.

2 Wavelet Analysis

For one dimensional wavelet analysis, a signal can be represented as a sum of wavelets at different time shifts and scales (frequencies) using discrete wavelet analysis (DWT). The DWT is capable of extracting the features of transient signals by separating signal components in both time and frequency. According to DWT, a time-varying function (signal) $f(t) \in L^2(R)$ can be expressed in terms of $\phi(t)$ and $\psi(t)$ as follows:

$$f(t) = \sum_k c_0(k)\phi(t-k) + \sum_k \sum_{j=1} d_j(k) 2^{\frac{-j}{2}} \psi(2^{-j}t - k)$$

$$= \sum_k c_{j0}(k) 2^{\frac{-j0}{2}} \phi(2^{-j0}t - k) + \sum_k \sum_{j=j0} d_j(k) 2^{\frac{-j}{2}} \psi(2^{-j}t - k)$$

where $\phi(t), \psi(t), c_0$, and d_j represent the scaling function, wavelet function, scaling coefficients at scale 0, and wavelet detailed coefficient at scale j, respectively. The variable k is the translation coefficient for the localization of a signal for time. The scales denote the different (high to low) frequency bands. The variable symbol j_0 is selected scale number.

The wavelet filter-banks approach was developed by Mallat [11]. We use Daubechies wavelet of order 7 (db7) [12] for wavelet analysis of mass spectrometry data and the boundary values are symmetrically padded. Multilevel discrete wavelet transform (DWT) is performed on mass spectrometry data. Given a mass spectrometry data ms of length N, the DWT consists of $Log_2 N$ levels at most. The first level produces, starting from ms, two sets of coefficients: approximation coefficients and detail coefficients. These vectors are obtained by convolving ms with the low-pass filter for approximation, and with the high-pass filter for detail, followed by dyadic decimation (downsampling). The next level splits the approximation coefficients in two parts using

the same scheme, and so on. Conversely, approximations and details are constructed inverting the decomposition step by inserting zeros and convolving the approximation and detail coefficients with the reconstruction filters. Figure 1 shows wavelet decomposition tree of mass spectra at 3 levels.

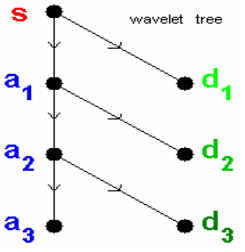


Fig. 1. Wavelet decomposition tree of mass spectra at 3 levels. Symbol s represents mass spectra; a_i, d_i represent approximation and detail at different levels respectively, where $i = 1,2,3$.

The purpose of detail coefficients is to detect the localized features in one of mass spectrometry data's derivatives based on multilevel wavelet decomposition. Detail coefficients determine the position of the change, the type of the change (a localized change in which derivative) and the amplitude of the change, using the compactness and finite energy characteristic of wavelet function. If the first levels of the decomposition can be used to eliminate a large part of the noise, the successive details lose progressively more high-frequency information. In our previous research [9], we have proved that the detail coefficients at 3rd level are efficient way to characterize the features of mass spectra. In this study we reconstruct the details based on detail coefficients at 3rd level. The corresponding m/z values are identified using selected GA features from detail at 3rd level.

The genetic algorithm is an evolutionary computing technique that can be used to solve problems efficiently for which there are many possible solutions [13]. We perform GA on wavelet features to select the best discriminant features and reduce dimensionality of wavelet feature space.

In our optimization problems, it is more natural to represent the variables directly as real numbers, which means there are no differences between the genotype (coding) and the phenotype (search space) [14]. A thorough review related to real-coded genetic algorithms can be seen in [15].

We define a chromosome C is a vector consisting of m gene variables $x_k, 1 \le k \le m$.

$$C = \{ (x_1,...,x_k,...,x_m) | 1 \le \forall i \le m : 1 \le x_i \le d_{\max}; 1 \le i, j \le m, i \ne j : x_i \ne x_j \}$$

Where $d_{\max}$ is the dimension number of wavelet reconstructed details. The algorithm creates initial population by ranking key features based on a two-way t-test with pooled variance estimate. We select different features in our study respectively to

evaluate the performance of classification. The algorithm creates a sequence of new populations based on different number of GA features. At each step, the algorithm uses the individuals in the current generation to create the next population.

3 Experiments and Results

In our research we use Prostate cancer dataset to evaluate the proposed algorithm. This data was collected using the H4 protein chip (JNCI Dataset 7-3-02) [16]. There are 322 samples: 190 samples with benign prostate hyperplasia with PSA levels greater than 4, 63 samples with no evidence of disease and PSA level less than 1, 26 samples with prostate cancer with PSA levels 4 through 10, and 43 samples with prostate cancer with PSA levels greater than 10. Each sample is composed of 15154 features. We combine benign prostate hyperplasia samples and those with no evidence of disease to form the normal class. The rest of the samples are classed into cancer category. We have 69 cancer samples and 253 normal samples. We use Correct Rate, Sensitivity, Specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV) and Balanced Correct Rate (BACC) to evaluate the performance of the proposed method [9,10].

The original prostate mass spectrometry data has 15154 features. The mass spectrometry data has high dimensionality and a lot of the information corresponds to m/z values that do not show any key changes during the experiment. Firstly we filter out mass spectrometry data with 60% variance and feature number with variance greater than 60^{th} percentile reduces to 6062. Then we perform wavelet analysis to extract the detail coefficients at 3^{rd} level, and obtain the reconstructed detail at 3^{rd} level based on orthogonal detail coefficients. Genetic algorithm is performed to select the best features from reconstructed detail at 3^{rd} level. Finally the significant m/z values are identified using the corresponding selected detail features.

K fold cross validation experiments [9, 10] are used to evaluate the performance of selected GA features. We run 10 times for each K fold cross validation experiments, where K=3, 4 and 5. We perform the GA on wavelet reconstructed details using different number of variables. Table 1 shows the performance of selected GA features from details at 3^{rd} level. When 18 GA features are selected, 97.07%, 97.20%, and 97.33% accuracy are obtained for 3, 4, and 5 fold cross validation experiments respectively. Figure 2 shows filtered mass spectra and 18 selected GA features from detail at 3^{rd} level. Figure 3 shows eighteen selected features from reconstructed details for test samples of Prostate cancer dataset based on 3 fold cross validation experiments. Figure 4 shows 97.07% performance of 18 key m/z values for test samples. 18 key m/z values are (29.506, 63.2295, 65.0231, 123.3446, 264.612, 270.7192, 305.5544, 384.8361, 412.0294, 505.8734, 509.6587, 843.4957, 1155.8994, 1570.9578, 6039.0984, 7120.7144, 8018.3359, 10113.674).

When 31 GA features are selected, 98.18%, 99.02%, and 99.19% accuracy are obtained for 3, 4, and 5 fold cross validation experiments respectively. Figure 5 shows 31 selected features from reconstructed details for test samples of Prostate cancer dataset based on 3 fold cross validation experiments. 31 key m/z values are (57.7154,

63.378, 110.0426, 118.2167, 124.3832, 125.6354, 126.2638, 264.612, 362.1142, 476.914, 503.3577, 506.2933, 540.0151, 1354.1112, 1417.3344, 1570.2181, 1615.6586, 1632.206, 1916.5929, 2354.1696, 2377.7745, 2417.0717, 3105.2114, 4070.0545, 5252.791, 6039.0984, 8475.9491, 8928.6153, 9057.8272, 9589.44, 10130.575).

Levner [7] performed 3 fold cross validation experiments using different methods. Their best BACC result is 90.6% using boosting based feature selection algorithm. When 18 GA features are selected, we obtain 97.07% accuracy and 97.53% BACC for 3 fold cross validation experiments. 98.18% accuracy and 97.79% BACC are achieved when 31 GA features are optimized. Our proposed method performs better than t-test, P-test, PCA/LDA, SFS, boosted method, etc.

Table 1. Performance of selected GA features from detail at 3rd level

GA features	K fold	Correct rate	Sensitivity	Specificity	PPV	NPV	BACC
	3	0.9707	0.9834	0.9673	0.8912	0.9954	0.9753
18	4	0.9720	0.9855	0.9684	0.8947	0.9959	0.9769
	5	0.9733	0.9855	0.9700	0.8995	0.9959	0.9777
	3	0.9818	0.9710	0.9848	0.9456	0.9920	0.9779
31	4	0.9902	0.9879	0.9908	0.9669	0.9967	0.9894
	5	0.9919	0.9913	0.9921	0.9716	0.9976	0.9917

This Table shows performance of low resolution prostate cancer dataset. PPV stands for Positive Predictive Value; NPV stands for Negative Predictive Value. BACC stands for Balanced Correct Rate.

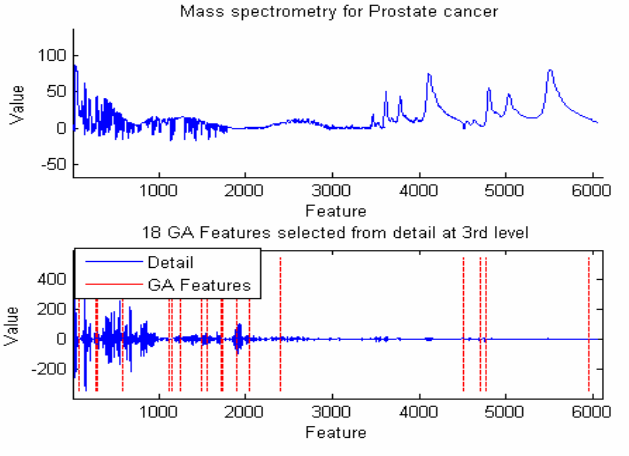


Fig. 2. Filtered mass spectra and 18 selected GA features from detail at 3rd level

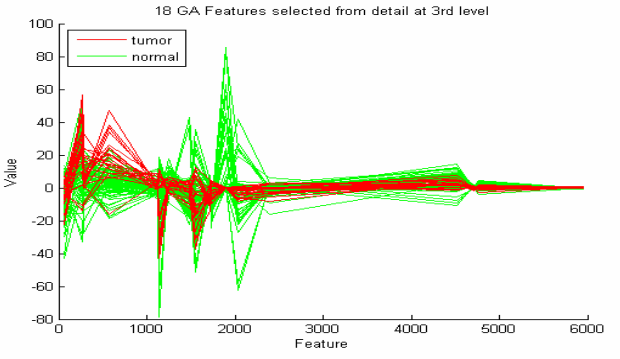


Fig. 3. Eighteen selected features from reconstructed details for test samples of Prostate cancer dataset based on 3 fold cross validation experiments

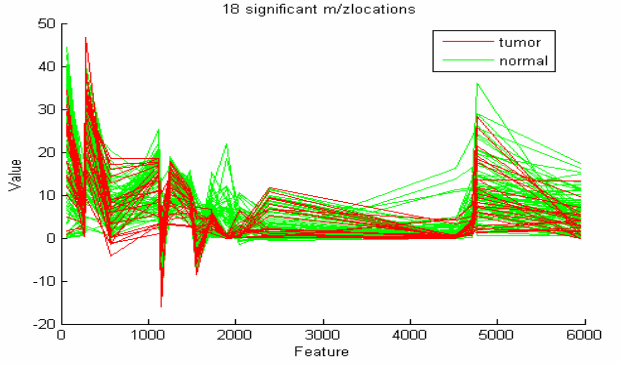


Fig. 4. 97.07% performance of 18 key m/z values for test samples of Prostate cancer dataset based on 3 fold cross validation experiments

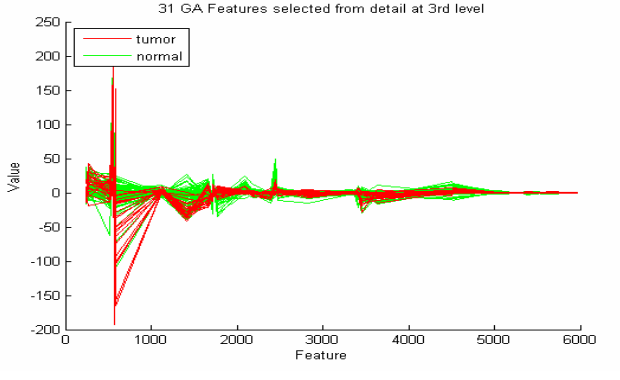


Fig. 5. Thirty one selected features from reconstructed detail for test samples of Prostate cancer dataset based on 3 fold cross validation experiments

4 Conclusions

In this paper we use multilevel wavelet analysis to extract the features of high dimensional mass spectrometry data. In order to find significant key m/z values of mass spectra, wavelet details based on the orthogonal wavelet detail coefficients at 3rd level are reconstructed, and characterize the features of mass spectra in original data space. Then we perform genetic algorithm on reconstructed detail to select the best features. Experiments are carried on Prostate cancer dataset. Good performance is achieved based on selected GA features. Finally the corresponding significant m/z values are identified based on the optimized wavelet detail features.

Acknowledgements. This work is supported by international collaboration project of Shandong Province Education Department, China.

References

1. Herrmann, P.C., Liotta, L.A., Petricoin, E.F.: Cancer Proteomics: The tate of the Art. Dis. Markers 17, 49–57 (2001)
2. Wright Jr., G.L., Cazares, L.H., Leung, S.M., Nasim, S., Adam, B.L., Yip, T.T., Schellhammer, P.F., Gong, L., Vlahou, A.: Protein Chip Surface Enhanced Laser Desorption/Ionization (SELDI) Mass Spectrometry: A Novel Protein Biochip Technology for Detection of Prostate Cancer Biomarkers in Complex Protein Mixtures. Prostate Cancer Prostatic Dis. 2, 264–276 (1999)
3. Vlahou, A., Schellhammer, P.F., Mendrinos, S., Patel, K., Kondylis, F.I., Gong, L., Nasim, S., Wright, J.: Development of a Novel Proteomic Approach for the Detection of Transitional Cell Carcinoma of the Bladder in Urine. Am. J. Pathol. 158, 1491–1520 (2001)
4. Lilien, R.H., Farid, H., Donald, B.R.: Probabilistic Disease Classification of Expression-Dependent Proteomic Data from Mass Spectrometry of Human Serum. Computational Biology 10 (2003)
5. Wu, B., Abbott, T., Fishman, D., McMurray, W., Mor, G., Stone, K., Ward, D., Williams, K., Zhao, H.: Comparison of Statistical Methods for Classification of Ovarian Cancer Using Mass Spectrometry Data. Bioinformatics 19 (2003)
6. Jeffries, N.O.: Performance of a Genetic Algorithm for Mass Spectrometry Proteomics. BMC Bioinformatics 5 (2004)
7. Levner, I.: Feature Selection and Nearest Centroid Classification for Protein Mass Spectrometry. BMC Bioinformatics 6 (2005)
8. Yu, J.S., Ongarello, S., Fiedler, R., Chen, X.W., Toffolo, G., Cobelli, C., Trajanoski, Z.: Ovarian Cancer Identification Based on dimensionality Reduction for High-throughput Mass Spectrometry Data. Bioinformatics 21, 2200–2209 (2005)
9. Liu, Y.: Feature Extraction for Mass Spectrometry Data. In: Li, K., Li, X., Irwin, G.W., He, G. (eds.) LSMS 2007. LNCS (LNBI), vol. 4689, pp. 188–196. Springer, Heidelberg (2007)
10. Liu, Y.: Cancer Classification Based on Mass Spectrometry. In: Masulli, F., Mitra, S., Pasi, G. (eds.) WILF 2007. LNCS (LNAI), vol. 4578, pp. 596–603. Springer, Heidelberg (2007)
11. Mallat, S.: A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 11, 674–693 (1989)

12. Daubechies, I.: Orthonormal Bases of Compactly Supported Wavelets. Communications on Pure and Applied Mathematics 41, 909–996 (1988)
13. Holland, J.H.: Adaptation in Natural and Artificial Systems. MIT Press, Cambridge (1992)
14. Blanco, M., Delgado, M.C.: Pegalajar: A Real-coded Genetic Algorithm for Training Recurrent Neural Networks. Neural Networks 14, 93–105 (2001)
15. Herrera, F., Lozano, M., Verdegay, J.L.: Tackling Real-coded Genetic Algorithms: Operators and Tools for Behavioral Analysis. Artif. Intell. Rev. 12, 265–319 (1998)
16. Petricoin, E.F., Ornstein, D.K., Paweletz, C.P., Ardekani, A., Hackett, P.S., Hitt, B.A., Velassco, A., Trucco, C., Wiegand, L., Wood, K., Simone, C.B., Levine, P.J., Linehan, W.M., Emmert-Buck, M.R., Steinberg, S.M., Kohn, E.C., Liotta, L.A.: Serum Proteomic Patterns for Detection of Prostate Cancer. J. Natl. Cancer Inst. 94, 1576–1578 (2002)

Protein Secondary Structure Prediction Based on Ramachandran Maps

Yen-Ru Chen, Sheng-Lung Peng*, and Yu-Wei Tsay

Department of Computer Science and Information Engineering
National Dong Hwa University
Hualien 974, Taiwan
slpeng@mail.ndhu.edu.tw

Abstract. High sequence similarity of two proteins does not imply the relation in their structural similarity. In this paper, we propose a protein secondary structure prediction method based on Ramachandran maps. According to the distribution of Ramachandran plot on ϕ and ψ backbone conformational angles, the series of backbone dihedral angles are regarded as a sequence of input states. Individual residues are classified into different secondary structures. In one of our results, protein hemoglobin 1A6M aligns 2FAM sequence sharing 11.74% similarity. However, 1A6M shares 81.82% of the common structure with 2FAM. Results display that our proposed method provides a promotive advantage on protein secondary structure prediction, without depending on sequence similarity.

1 Introduction

Secondary structure is formally defined by the hydrogen bonds of the biopolymer, as observed in an atomic-resolution structure. Its prediction from amino acid sequence can provide information to build the initial model for molecular simulation, it also serve as important building blocks in comparative modeling. The precision of protein secondary structure prediction greatly affects the quality of generated templates. Therefore, a protein secondary structure prediction method from multiple aligned homologous sequences is presented with an overall per residue three-state accuracy of 70% [6]. Theoretically, if a protein sequence is given, then its senior structures can be determined by various simulation methods. However, a possible drawback of this approach is strongly biased by the content of the protein sequence similarity. If similarity between the target sequence and the template sequence is detected, structural similarity can be assumed. In general, 30% sequence identity and some known protein structures are required to generate a useful model.

For this reason, a method is proposed to construct protein secondary structure without adapting amino acid sequence. Distinct Ramachandran regions suggest individual mapping for the formation of predicted protein secondary structure. We denominate it as *Ramachandran Adjacency Index* for protein secondary structure predication, abbreviated as (RAI). Once the corresponding

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 204–211, 2008.
© Springer-Verlag Berlin Heidelberg 2008

map of Ramachandran is determined, the coordinates can be transferred to a protein secondary structure according to the limitation of atoms' moving. Empirical distribution of dihedral plane indicates the relation between amino acids (PROSS) [14]. Therefore, the protein secondary structure can be obtained.

In this paper, we present a hybrid approach to predict secondary structure of a unknown structure protein. The steps are as follows. The first step is to produce a set of Ramachandran axis from unknown structure protein. Then, the second step assigns the corresponding expected secondary structure to the unknown protein from RAI index training model. Finally, each simulated Ramachandran region is allocated by a particular structure.

1.1 Preliminaries

In a polypeptide, the main chain N-terminal and C-terminal bonds relatively are free to rotate. These rotations are represented by the torsion angles phi and psi, respectively. A Ramachandran plot is developed by Gopalasamudram Narayana Ramachandran in 1968 [14]; meanwhile, it is a way to visualize dihedral angles ϕ and ψ of amino acid residues in protein structure. Ramachandran plot shows the possible conformations of ϕ and ψ angles for a polypeptide chain. In a Ramachandran plot, the core or allowed regions are the areas in the plot showing the preferred regions for ϕ and ψ angle pairs for residues in a protein. Presumably, if the determination of protein structure is reliable, most pairs will be in the favored regions of the plot and only a few will be in disallowed regions. There are various definitions of the allowed areas in Ramachandran plots.

1.2 Related Works

The accuracy of secondary structure prediction methods has been improved significantly by the use of aligning protein sequences. [12]. Besides, a two-stage neural network has been used to predict protein secondary structure based on the position specific scoring matrices generated by PSI-BLAST [5]. And by adding neural network units that detect periodization in the input sequence, the use of tertiary structural class causes a marked increase in accuracy. The best case for prediction is 79% for the class of all-alpha proteins [7]. A method has also been developed to check against from an updated protein database. This method correctly predicts 69% of amino acids for a three-state description of the secondary structure, including helix, 3 sheet, and coiled in the whole database [2].

Meanwhile, the effect of training a neural network with different types of multiple sequence alignment profiles derived from the same sequences is shown to provide a range of accuracy from 70.5% to 76.4% [1]. Another method is using evolutionary information contained in multiple sequence alignments as input to neural networks. It has a sustained overall accuracy of 71.6% in a multiple cross-validation test on 126 unique protein chains [11]. The above methods are composed by several neural network approaches and taking into account for sequence similarity. In addition, there is a notable algorithm called PHD based

on a neural network learning process system [10]. This is the first method to break the 65% boundary on Q3 accuracies of secondary structure prediction methods. Besides, a method has been tested on 59 proteins in the database and yields 72% success in class prediction, 61.3% of residues correctly predicted also for three states [3].

2 Our Method

In this section, a method is proposed to predict protein secondary structure without concerning its sequence similarity. This approach generate a series of ϕ and ψ angles for a given protein sequence, according to the RAI-index. As long as the angles of given protein has been resolved, a secondary information can be decided by its allowable Ramachandran region. Hence, the modeled protein secondary structure can be treated as a referral for 3D protein conformation.

2.1 Ramachandran Adjacency Index

The influence of protein nearest neighbors for the angles on ϕ and ψ is evaluated as interrelated, so we make an assumption that the ϕ and ψ angles of n^{th} amino acid are infected by the pentapeptide as $(n-2)$, $(n-1)$, $(n+1)$, and $(n+2)$. In [13], it tried to use tripeptide and their corresponding ϕ and ψ angles to predict the unknown protein structure. Besides, it also mentioned that the formation of a helix including position (n) obviously may involve amino acids at positions $(n-4)$, $(n-3)$, $(n+3)$, and $(n+4)$. Consequently, by considering the problem on computational complexity and data set size, the pentapeptide relation is chosen as our rule, instead of using octpeptide. The training algorithm of RAI-index construction is given as follows.

Input: A dataset with amino acids and their ϕ, ψ-angles at each position.
Output: Combinations of pentapeptide and its ϕ, ψ-angles.

```
1: Table initializing with structures AAseq (input protein sequence), phi (angle
   of φ), psi (angle of ψ), index, RT (repeat times), and PN (protein number)
2: for p = 1 to PN do
3:    ln = the length of the protein
4:    for i = 3 to (ln − 2) do
5:       query the AAseq residue S of (i−2)(i−1)(i)(i+1)(i+2) in RAI-Index
6:       if YES then
7:          store the value of φ and ψ angles of responding AAseq
8:          RT[AAseq]=RT[AAseq]+1
9:       else
10:         insert the combination of pentapeptide and record it in RAI-Index
11: L= Length of of RAI-Index
12: for j=1 to L do
13:    calculate the average of φ and ψ
14: record the RAI-Index
```

Note that ϕ and ψ data on proteins of known 3D structures are available from dataset. Thus, the angles can be computed by RAI-index. According to this algorithm, training data set is updated constantly for every pentapeptide. Therefore, the ϕ and ψ angles will become more accurate unless the data is incorrect. In order to prevent the noise occurred from uncertain peptide during the training procedure, a threshold is given to adjust RAI-index stably. If one of the pentapeptide distances is larger than the threshold, then it would not be adapted during the training procedure.

2.2 Secondary Structure Transformation

Once the ϕ and ψ data has been trained to establish RAI-index by pentapeptide, it is easy to return a series of ϕ and ψ angles if a queried sequence is similar to RAI-index. According to Fitzkee's results [4], the whole Ramachandran plot has been classified into five specific zones. They are α-helix, α'-helix, β-strand, Polyproline helix, and κ. Each of them is exactly represented as one kind of secondary structure. They are summarized in Table 1.

Table 1. The region represented for each secondary structure

Secondary Structure	ϕ range	ψ range
α-helix	A circle of radius 7° centered about (−63°,−45°)	
α'-helix	$-75.0° \leq \phi \leq -45.0°$	$-60.0° \leq \psi \leq -30.0°$
β-strand	$-135.0° \leq \phi \leq -105.0°$	$120.0° \leq \psi \leq 150.0°$
Polyproline helix	$-80.0° \leq \phi \leq -55.0°$	$130.0° \leq \phi \leq 155.0°$
κ	All other accessible regions	

In general, this approach is implemented by PROSS [14]. The difference between α-helix and α'-helix is the amounts of points. The α'-helix covers α-helix. In this formula, α'-helix is applied to determine whether it is helix or not. With the objective of stabilizing protein conformation, a new limitation rule is added to our proposed algorithm. That is, the β conformation cannot be followed by the α conformation. The following example demonstrates the usage of each secondary region.

Suppose an amino acid sequence is given as follows: Ile-Ile-Cys-Thr-Ile-Gly-Pro-Ala-Ser-Arg-Ser-Val. The primary sequence is looked up for each 5-residue (*i.e.*, a pentapeptide). The first pentapeptide —Ile-Ile-Cys is certainly failed during the process of inquiring RAI-Index. It misses the former two amino acids. Similarly, the last pentapeptide misses the latter two amino acids. In this case, we provide another method to solve these special cases in Section 2.3. In the following, we suppose each pentapeptide can be entirely matched in RAI-Index. That is, it returns a Ramachandran value.

Once the Ramachandran value of the protein is determined, each residue can be mapped to a specific ϕ and ψ ranges by referring Table 1. After the conversion

Table 2. The corresponding angles of Φ and Ψ for given protein sequence. **PR** indicates its PROSS region, and **SS** represents each secondary structure. **H**:helix, **E**:strand, **P**:PP2, **T**:turn, **C**:coil.

AA	Ile	Ile	Cys	Thr	Ile	Gly	Pro	Ala	Ser	Arg	Ser	Val
ϕ	-124.2	-95.0	-111.2	-91.7	-74.8	-147.0	-55.7	-86.4	-135.8	-81.5	-64.4	-47.0
ψ	111.2	119.8	136.7	104.3	138.8	-164.6	-28.8	-12.0	25.9	-27.2	126.8	-33.2
PR	Ci	Di	Ck	Di	Ek	Ba	Ef	Df	Bg	Df	Ek	Ee
SS	E	E	E	E	P	C	T	T	C	C	P	H

for each corresponding structure, the amino acid sequence will be transformed into a secondary form, *e.g.*, EEEEPCTTCCPH. Accordingly, **H** is for α-helix, **E** is for β-strand, **P** is for Polyproline, **T** is for β-turn, and **C** is for coil.

2.3 Estimation for Uncertain RAI-Index

Since the RAI-index table has been established, a simulated Ramachandran plot can be obtained from a query sequence. However, there is a problem if there is a mismatch in looking up RAI-index during the transformation. For example, suppose the value of fragment 'Val-Leu-Ser-Glu-Gly' cannot be found in RAI-index, indicating the ϕ and ψ angles of this fragment is empty in database. Therefore, it is necessary to estimate a possible value by its chemical property.

In the study of [13], if the segment $(n-1)$ (n) $(n+1)$ is missed in RAI-index, then $(n-1)(X)(n+1)$, $(X)(n)(n+1)$, and $(n-1)(n)(X)$ can be used to replace the missing RAI-index, where X can be any one of the 20 amino acids. The concept of [13] tried to use the nearest residue replacing the failed one. It is easy to observe that the calculation of pentapeptide is more complicated than tripeptide since the possibility of missing segments are large. For this reason, $(X)(n-1)(n)(n+1)(n+2)$, $(n-2)(X)(n)(n+1)(n+2)$, $(n-2)(n-1)(X)(n+1)(n+2)$,$(n-2)(n-1)(n)(X)(n+2)$, and $(n-2)(n-1)(n)(n+1)(X)$ are used to refer the missing RAI-index.

We give another example for the case of empty RAI-Index as follows. Suppose residues **Gly-Cys-Asp-Asn-Glu** is not in the data set of RAI-index. In this case, this segment should refer to a similar pattern in RAI-index that is non-empty. Next, the RAI-Index searches all data set to replace each peptide candidates and record it. RAI-Index returns the outcome that can be substituted by each peptide. According to the polarity of their matching peptide, the segment of **Gln-Cys-Asp-Asn-Glu** should be the proper candidate in RAI-index, since Aspartic Acid (Asp) and Glutamine (Gln) are regarded as the same polarity group.

3 Results

In our results, 220 proteins of myoglobin are regarded as training data set, 10 proteins of other myoglobin are treated as test data set. The criteria of predicted

Table 3. The similarity of proteins with 1A6M (Myoglobin)

Protein ID	Similarity(%)		Annotation
	Structure	AA similarity	
2FAM	81.82	11.74	Ferric Complex With Thiocyanate
3MBA	81.82	11.74	Myoglobin (Fluoride Complex)
2FAL	79.55	11.74	Myoglobin (Ferric) Complex With Cyanide
1DM1	75.00	13.66	Crystal Structure Of H(E7)
1MYT	75.00	35.53	Myoglobin (Met)
1L2K	73.85	98.38	Neutron Structure Determination
1A6K	73.28	100	Aquomet-Myoglobin, Atomic Resolution
1A6N	72.52	100	Deoxy-Myoglobin, Atomic Resolution
1DWR	68.18	88.46	Wild-Type Complexed With Co.
2IN4	60.90	88.46	Charge Neutralized Heme
1DXC	55.15	94.84	Co Complex Of Myoglobin Mb-Yqr At 100k
1EBT	48.48	4.86	Met-Myoglobin:cyanide Complex

structure with known structure are mainly depending on global alignment. Although myoglobin performs as a temporary store for the oxygen, it can be observed that some of their conformations are various. A partial results of our proposed rule display a great discrimination between sequence and structure in Table 3. Indeed, high protein sequence similarity implies a great probability of structural similarity. However, this principle is not suitable for all cases in protein folding. That is, an alternative approach is required to solve this problem. If 2FAM is aligned in the data set, it cannot describe its secondary structure presumably. The alignment of protein 1A6M with 2FAM scores only 11.74%; in contract, 1A6M shares 81.82% of the common secondary structures with 2FAM. And also, protein 1DM1 shares 75% structural similarity with 1A6M, but aligning only 13.66% sequence similarity with it. Figure 1. shows that myoglobin 1A6M is actually similar to 2FAM, 1DM1, and 1MYT in their 3D structures. All of their sequence similarities are low, instead of their structural conformation.

4 Discussion

In this paper, we propose a reasonable and reliable method to predict protein secondary structures. Currently, the analysis of protein sequence and structure relationships focuses on their similarity and structural similar proteins. That is, gene functionality determines the similarity of protein sequences. However, the relation between protein sequence and structure is not always coordinative. Sequence homology does not completely confirm the partial sequence similarity

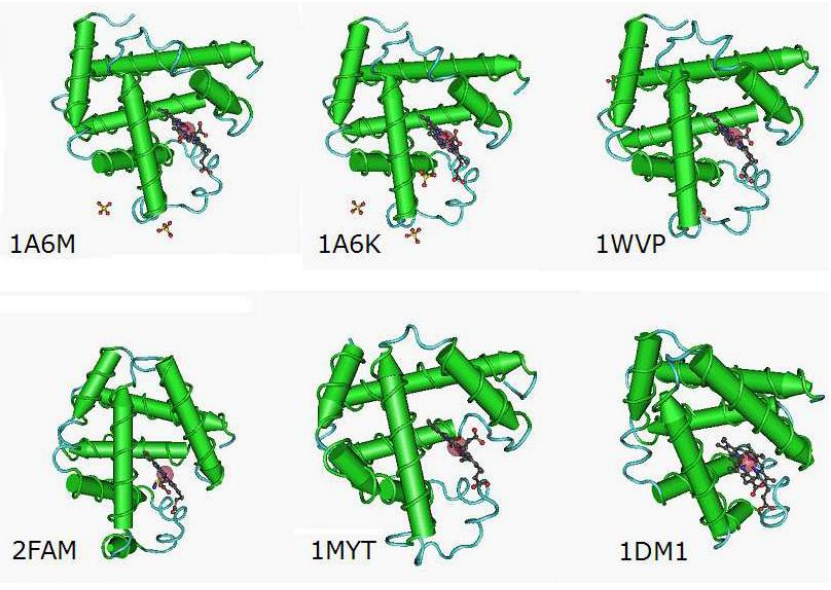


Fig. 1. The results of myoglobin 1A6M and other proteins which are dissimilar in their protein sequences

that may conserve their structural homologous. For these reasons, our method expects to present a measure of protein structural similarity, based on a simple approach by allowable Ramachandran region transformation. Hence, one of our future works will focus on macromolecular polymers and attempt to quantify the relation between protein secondary structure and its conformation. Also, the estimation of uncertain RAI-Index can be raised to two missing residues. Finally, a comparison with other secondary structure methods will be discussed in future.

References

1. Cuff, J.A., Barton, G.J., Langridge, R.: SOPM: A Self-Optimized Method for Protein Secondary Structure prediction. Proteins 40, 502–511 (2000)
2. Deléage, G., Geourjon, C., Langridge, R.: Application of Multiple Sequence Alignment Profiles to Improve Protein Secondary Structure Prediction. Protein Engineering 7, 157–164 (1994)
3. Deléage, G., Roux, B.: An Algorithm for Protein Secondary Structure Prediction Based on Class Prediction. Protein Engineering 1, 289–294 (1987)
4. Fitzkee, N.C., Rose, G.D.: Steric Restrictions in Protein Folding: An R-helix Cannot Be Followed by A Continguous β-strand. Protein Sci. 13, 633–639 (2004)
5. Jones, T.D.: Protein Secondary Structure Prediction Based on Position-Specific Scoring Matrices. J. Mol. Biol. 292, 195–202 (1999)

6. King, R.D., Sternberg, M.J.: Identification and Application of the Concepts Important for Accurate and Reliable Protein Secondary Structure Prediction. Protein Sci. 5, 2298–2310 (1996)
7. Kneller, D.G., Cohen, F.E., Langridge, R.: Improvements in Protein Secondary Structure Prediction by an Enhanced Neural Network. J. Mol. Biol. 214, 171–182 (1990)
8. Murzin, A.G., Brenner, S.E., Hubbard, T., Chothia, C.: SCOP: A Structural Classification of Proteins Database for the Investigation of Sequences and Structures. J. Mol. Biol. 247, 536–540 (1995)
9. Ordway, G.A., Garry, D.J.: Myoglobin: An Essential Hemoprotein in Striated Muscle. J. Exp. Biol. 207, 3441–3446 (2004)
10. Rost, B., Sander, C.: Prediction of Protein Secondary Structure at Better Than 70% Accuracy. J. Mol. Biol. 232, 584–599 (1993)
11. Rost, B., Sander, C., Langridge, R.: SOPM: a Self-Optimized Method for Protein Secondary Structure Prediction. Proteins 19, 55–72 (2004)
12. Salamov, A.A., Solovyev, V.V.: Protein Secondary Structure Prediction using Local Alignments. J. Mol. Biol. 268, 31–36 (1997)
13. Wu, T.T., Kabat, E.A.: An Attempt to Evaluate the Influence of Neighboring Amino Acids $(n-1)$ and $(n+1)$ on the Backbone Conformation of Amino Acid (n) in Proteins. Use in Predicting the Three-Dimensional Structure of the Polypeptide Backbone of Other Proteins. J. Mol. Biol. 75, 13–31 (1973)
14. PROSS: Dihedral Angle-Based Secondary Structure Assignment, `http://www.roselab.jhu.edu/utils/pross.html`

Probe Selection with Fault Tolerance

Sheng-Lung Peng[1,⋆], Yu-Wei Tsay[1], Tai-Chun Wang[1], and Chuan Yi Tang[2]

[1] Department of Computer Science and Information Engineering
National Dong Hwa University, Hualien 974, Taiwan
[2] Department of Computer Science
National Tsing Hua University, Hsinchu 300, Taiwan
slpeng@mail.ndhu.edu.tw.

Abstract. Microarray techniques play an important role for testing some reactions of diseases which are caused by viruses. Probes in microarray are one kind of the most important materials. Usually, scientists use a unique probe for marking a special target sequence. Thus, for identifying n different viruses, we need n different probes. Recently, some researchers study non-unique probes to identify viruses by using less number of probes. In this case, a virus can be identified by a combination of some probes. In this paper, we study the problem of finding a set of probes that can identify all the given targets. We consider the k-fault tolerance selection of probes. That is, if any k probes fail, then we still can identify each target. We propose a practical algorithm for this k-fault tolerance probe selection problem. Some experiments are studied on SARS, H5N1, and so on.

1 Introduction

Research for the human nosography and immunology is important today [1]. SARS, AIV, cervical cancer, acquired immunodeficiency syndrome (AIDS), and so on, are all infected from viruses. Many researchers use microarray for clinical experiments [4,8]. A microarray usually uses parallel analysis for an expression of thousands of genes. Because it can be formed by different types (*e.g.*, oligonucleotide and cDNA), the applications of microarray are very wide [5].

Recent works or softwares try to find unique probes for the target sequences [2,11,12]. That is, each probe can only hybridize with one target sequence. In this case, for n different targets, we need n different probes to identify these targets. If we allow one probe fault, then for the premise of unique sequences, each target will need two unique sequences as probes. It is not hard to see that the number of probes will be doubled. It is very impractical. Thus, some researchers focus on non-unique probes [9,10]. Since each of non-unique probes can hybridize with more than one target, fewer probes are needed. In this case, a target is identified by a combination of selected probes.

Klau *et al.* proposed an integer linear programming algorithm to find a minimum set of non-unique probes which can identify all the input targets [6,7].

⋆ Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 212–219, 2008.
© Springer-Verlag Berlin Heidelberg 2008

In this paper, we propose a simple greedy algorithm to find a minimal set of non-unique probes for recognizing the input targets.

2 Preliminaries

A *probe* is a single-strand DNA or RNA molecule with a specific sequence labeled either radioactively or immunologically. It is used to detect the complementary sequence by hybridization for identifying unknown species or gene functions.

When designing probes, we need to get a single-strand DNA or RNA sequence. Then we use primers to help us produce the complement subsequences of the ss-DNA (RNA). Each of them is between 25 and 80 bp. To obtain probes, we need to consider the temperature, GC contents rate, DNA secondary structure, and cross-hybridization. Most known probes are stored in NCBI GeneBank. There are several softwares can help us to find probes, such as MAPS [2], GoArray [11], OligoArray [12], and so on.

Figure 1 shows an example. For unique probes, each target needs one unique probe to identify itself. In the non-unique probes, we use less probes to identify all the targets.

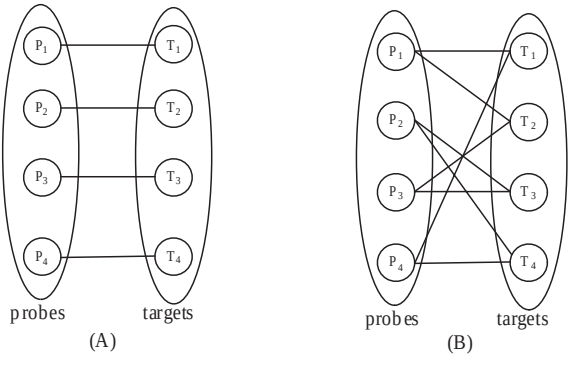


Fig. 1. Unique and non-unique probes

Table 1 shows a *relation table* according to Figure 1. In this table, if target T_i can hybridize to probe P_j, then the (i, j) entry is 1. According to this table, it is clear that we need four probes to identify all the targets by using unique probes. But for nun-unique probes, two probes, namely, P_1 and P_3, are sufficient. That is an advantage of using non-unique probes. Furthermore, in large transcript families, such as alternative splice variants of a gene, or in a large family of closely homologous genes, it is often impossible to find a unique probe of 25 bp as a signature for a specific variant [10]. So we consider about non-unique probes.

A *microarray* is an array containing thousands of probes. These probes are printed on a solid surface. Sometimes it is a glass or nylon membrane.

The sensitive of a microarray will be influenced by the reaction of target's concentration, temperature, and reaction time. Moreover, we also need to compute

Table 1. The relation matrixes for Figure 1 (A) and (B) respectively

	P_1	P_2	P_3	P_4
T_1	1	0	0	0
T_2	0	1	0	0
T_3	0	0	1	0
T_4	0	0	0	1

(A)

	P_1	P_2	P_3	P_4
T_1	1	0	0	1
T_2	1	0	1	0
T_3	0	1	1	0
T_4	0	1	0	1

(B)

the diffusion coefficient, adhesive force coefficient, the uniform level of reaction targets, and so on. Some researches show that a long oligonucleotide is more sensitive than short one ([3]). However, the design problem is very complicated. Chung *et al.* try to use a hierarchical method to design probes [3]. For exon chip, the length of the probes on the chip can be different. But for GeneChip, its length has the limitation of 25 bp.

Sometimes probes will fail because of the pollution during producing a microarray. So we need to add extra probes in order to make sure if some probe fails, then there still has another probe can take care of it. Table 2 shows the case that if any probe fails, then the probe set still can identify the four targets.

Table 2. Probe sets with 1-fault tolerance for Figure 1 (A) and (B) respectively

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
T_1	1	1	0	0	0	0	0	0
T_2	0	0	1	1	0	0	0	0
T_3	0	0	0	0	1	1	0	0
T_4	0	0	0	0	0	0	1	1

(A)

	P_1	P_2	P_3	P_4
T_1	1	0	0	1
T_2	1	0	1	0
T_3	0	1	1	0
T_4	0	1	0	1

(B)

3 Methods

In this section, we will introduce our greedy algorithm for finding a set of non-unique probes with k-fault tolerance.

3.1 Data Preprocessing

In our method, we use the software **Promide** proposed by Rahmann to find all the possible probes [9]. The input file is in FASTA format. After finding the

Table 3. An example of target-probe relation table

	P_1	P_2	P_3	P_4	P_5	P_6
T_1	1	0	0	1	1	0
T_2	0	1	0	0	1	1
T_3	1	1	1	1	1	1
T_4	0	1	1	1	0	0

Table 4. A distinguishable table obtained from Table 3

Target pairs	(T_1, T_2)	(T_1, T_3)	(T_1, T_4)	(T_2, T_3)	(T_2, T_4)	(T_3, T_4)
P_1	1	0	1	1	0	1
P_2	1	1	1	0	0	0
P_3	0	1	1	1	1	0
P_4	0	0	0	1	1	1
P_5	0	0	1	0	1	1
P_6	1	1	0	0	1	1

probes, we construct a relation table between probes and targets. Table 3 shows an example. For each target, if a probe can bind the target, we then assign the value 1 in the corresponding entry. Thus, we know which probe can affect which target. For distinguishing the targets, we translate the relation table into the *distinguishable table* (see Table 4). In the case that different probes have the same distinguishable ability, we only keep one. In Table 4, T_1 can hybridize to probe P_1, but T_2 cannot. That is, P_1 can distinguish T_1 and T_2. Thus, we record the value 1 in the (P_1-(T_1, T_2)) entry.

3.2 Relation Graph

Take Table 4 as an example. In order to illustrate the relationship between probes and target pairs, we use the distinguishable table to draw a *relation graph* (see Figure 2). In this graph, each edge denotes that the probe can differentiate the two targets in the target pair. By using non-unique probes, each target pair has more than two incident edges. That is, if we remove any one node from the probe set, any node in the target pairs still has at least one incident edge. It means that each target pair still can be distinguished by one probe. Based on this fact, it is not hard to see that a probe set is k-fault tolerance if every target pair in the graph has at least $k + 1$ edges. Figure 3 shows an example of 1-fault tolerance. In this graph, every target pair has at least two edges and any failure probe will not affect the distinguishable ability.

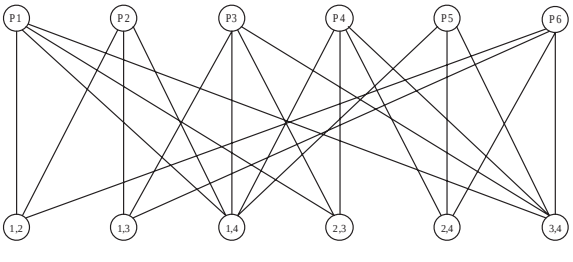


Fig. 2. The relation graph of Table 4

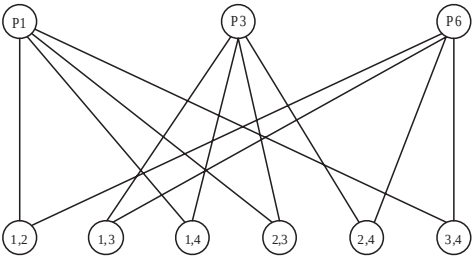


Fig. 3. An induced subgraph of Figure 2

Table 5. The main variables used in our algorithm

Input variables	
$relationArray$	It is a $n \times m$ matrix, where n is the number of possible probes and m is the number of target pairs.
$ftValue$	The number of fault tolerance.
Output variables	
$coverArray$	It is an array of size n. Each entry stores the number of target pairs that can be distinguished by the corresponding probe.
$ftArray$	It is an array of size m. Initially, every entry is 0. If every entry is at least $k+1$, then the found probe set is k-fault tolerance.
$identifyArray$	It is an array of size m. Each entry stores the number of probes that can be used to identify the target pair.
$ProbeSet$	A probe set outputted by our algorithm.

3.3 Probe Selection Algorithm

After constructing the relation graph, we start to find a probe set. We give some definition first for the variables which are used in our algorithm in Table 5.

Our algorithm is shown in Algorithm 1. It can be divided into three parts. The first part is from step 7 to step 10. This part helps us to find the smallest number in the $identifyArray$. By our experience, if the target pair that can be distinguished by few probes, then considering it first will obtain a good solution.

After finding a candidate target pair in the first part, we then go to the second part from step 13 to step 16. In this part, we find the largest number in the $coverArray$. That is, we select the probe that can distinguish the maximum number of target pairs.

The final part of our algorithm is to update all the related information after we select a probe. If the requirement of the fault tolerance does not fulfill, we then repeat the process again until we fulfill the requirement.

Algorithm 1. ProbeSelect($relationArray$, n, m, $ftValue$)

```
1: ftArray[i] = 0 for all i ∈ {1, ..., m}
2: ProbeSet = φ
3: loop = .T.
4: while loop do
5:    temp = ∞
6:    targetPair = 0
7:    for i = 1 to m do
8:       if ftArray[i] < ftValue and identifyArray[i] < temp then
9:          targetPair = i
10:         temp = identifyArray[i]
11:   temp = 0
12:   probe=0
13:   for i = 1 to n do
14:      if relationArray[i][targetPair] = 1 and temp < coverArray[i] then
15:         probe = i
16:         temp = coverArray[i]
17:   ProbeSet = ProbeSet ∪ {probe}
18:   loop = .F.
19:   for i = 1 to m do
20:      if relationArray[probe][i] = 1 then
21:         identifyArray[i] = identifyArray[i] − 1
22:         ftArray[i] = ftArray[i] + 1
23:      if ftArray[i] < ftValue then
24:         loop = .T.
25: return ProbeSet
```

The algorithm first finds the minimum number of $identifyArray$ (*e.g.*, $identifyArray[1]$). Then it finds the maximum number in $coverArray$ for target pair 1. Since probe 1 and probe 6 have the same value, we choose probe 1. After choosing probe 1, we update the remaining information. The selected probes are shown in Table 6, *i.e.*, $\{P_1, P_3, P_6\}$. It is not hard to check that any failure of selected probes does not loss the distinguishing ability for the input targets.

Table 6. The final result of our algorithm

(T_1, T_2)	(T_1, T_3)	(T_1, T_4)	(T_2, T_3)	(T_2, T_4)	(T_3, T_4)
P_1	P_6	P_1	P_1	P_6	P_1
P_6	P_3	P_3	P_3	P_3	P_6

4 Results

In this section, we show the result by using our algorithm to find a set of non-unique probes for some viruses (*e.g.*, SARS, HIV, avian's influenza A virus, and so on) with the ability of k-fault tolerance. Our virus data were downloaded from

Table 7. An experimental result of our algorithm

Dataset	No. of targets	No. of probes	0-FT selected	1-FT selected	2-FT selected	4-FT selected
SARS	121	1041	75	140	207	336
H5N1	136	1156	79	142	222	387
AIV	484	4356	216	252	730	1203
Retroviruses	54	10169	53	105	157	262

NCBI Database (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). The avian's influenza A virus was downloaded from the Influenza Virus Project web site which is maintained by NCBI. All the data are stored in FASTA format.

In the experiment of AIV's data, we mix all the H5N1's PB2 sequences and other PB2 sequences of different subtypes of avian's influenza A virus. Table 7 shows the result. The number k in k-FT (k from 0 to 4) is the fault tolerant number.

In our experiment, if the data are highly conservative such as SARS, then it is easy for **Promide** to find many possible non-unique probes. On the other hand, if the data are not highly conservative, **Promide** will find a lot of unique probes. In this case, our algorithm obtains a lot of probes like Retroviruses.

5 Discussion

The production of microarrays is costly. An important factor is the number of probes used in the microarray. Besides, the precision of tests for microarrays is another factor. These motivate this research. By the experimental result, our method shows that we can reduce the number of probes used for identifying several groups of viruses. Further, our method can be easily extended to k-fault tolerance, *i.e.*, if there are k failure probes, the remaining probes can still be used to recognize the input viruses.

In fact, our probe selection algorithm is an approximation algorithm. If the input size is not too large, we may use a branch-and-bound algorithm to find an optimal solution. Second, for the tool **Promide**, it is unable in dealing with an oversize sequences. Thus, it is necessary to find or design a new tool for large scale input size. Finally, while the cost of building microarrays is down, fault tolerance becomes an important issue. Our method obtains only an approximated result. Designing new algorithms considering fault tolerance will be a new direction.

References

1. Abbas, A.K., Lichtman, A.: Cellular and Molecular Immunology, 5^{th} edn. Saunders, Philadelphia (2005)
2. Bushel, P.R., Hamadeh, H., Bennett, L., Sieber, S., Martin, K., Nuwaysir, E.F., Johnson, K., Reynolds, K., Paules, R.S., Afshari, C.A.: Maps: a microarray project system for gene expression experiment information and data validation. Bioinformatics 17, 564–565 (2001)

3. Chung, W.H., Rhee, S.K., Wan, X.F., Bae, J.W., Quan, Z.X., Park, Y.H.: Design of long oligonucleotide probes for functional gene detection in a microbial community. Bioinformatics 21, 4092–4100 (2005)
4. Cooper, C.S.: Applications of microarray technology in breast cancer research. The Institute of Cancer Research 3, 158–175 (2001)
5. He, Y.D., Dai, H., Schadt, E.E., Cavet, G., Edwards, S.W., Stepaniants, S.B., Kleinhanz, S.D.R., Jones, A.R., Shoemaker, D.D., Stoughton, R.B.: Microarray standard data set and figures of merit for comparing data processing methods and experiment designs. Bioinformatics 19, 956–965 (2003)
6. Klau, G.W., Rahmann, S., Schliep, A., Vingron, M., Reinert, K.: Optimal robust non-unique probe selection using integer linear programming. Bioinformatics 20(Suppl. 1), i186– i193 (2004)
7. Klau, G.W., Rahmann, S., Schliep, A., Vingron, M., Reinert, K.: Integer linear programming approaches for non-unique probe selection. Discrete Applied Mathematics 155, 840–856 (2007)
8. Petricoin, E.F., Hackett, J.L., Lesko, L.J., Puri, R.K., Gutman, S.I., Chumakov, K., Woodcock, J., Feigal Jr., D.W., Zoon, K.C., Sistare, F.D.: Medical applications of microarray technologies: a regulatory science perspective. Nature Genetics Supplement 32, 474–479 (2002)
9. Rahmann, S.: Rapid large-scale oligonucleotide selection for microarrays. In: The First IEEE Computer Society Bioinformatics Conference, pp. 54–63 (2002)
10. Rahmann, S., Muller, T., Vingron, M.: Non-unique probe selection by matrix condition optimization. Currents in Computational Molecular Biology (2004)
11. Rimour, S., Hill, D., Militon, C., Peyret, P.: Goarrays: highly dynamic and efficient microarray probe design. Bioinformatics 21, 1094–1106 (2005)
12. Rouillard, J.M., Zuker, M., Gulari, E.: Oligoarray 2.0: design of oligonucleotide probes for dna microarrays using a thermodynamic approach. Nucleic Acids Research 31, 3057–3062 (2003)

Image Reconstruction Using a Modified Sparse Coding Technique*

Li Shang

Department of Electronic Information Engineering, Suzhou Vocational University, Suzhou, Jiangsu 215104, China
sl0930@jssvc.edu.cn

Abstract. A novel image reconstruction method for natural images using a modified sparse coding (SC) algorithm is proposed by us. This SC algorithm exploited the maximum Kurtosis as the maximizing sparse measure criterion, and a fixed variance term of sparse coefficients is used to yield a fixed information capacity. The experimental results show that using our algorithm, the natural images' feature basis vectors can be successfully extracted. Furthermore, compared with the standard SC method, the experimental results show that our algorithm is indeed efficient and effective in performing image reconstruction task.

Keywords: Sparse coding, Kurtosis, Fixed variance, Feature extraction, Image reconstruction.

1 Introduction

In recent years, the problem how to find optimal algorithm for efficient coding and redundancy reduction has been the subject considered by more and more people. As we know, recent techniques such as principal components analysis (PCA) [1], independent components analysis (ICA) [2], sparse coding (SC) [3], and non-negative matrix factorization (NMF) [4] have been used in the area of image coding successfully by minimizing the dependencies between elements of the code under certain appropriate constraints. These above algorithms are typically data-adaptive representations, which are fitted to the statistics of the data and can be learned directly from the observed data.

Standard sparse coding (SC) algorithm was early proposed by Olshausen and Field in 1996 [3]. More over, when the method of independent component analysis (ICA) is applied to natural data, ICA is equivalent to SC [5]. However, ICA emphasizes independence over sparsity in the output coefficients, while SC requires that the output coefficients must be sparse and as independent as possible. Because of the sparse structures of natural images, SC is more suitable to process natural images than ICA. Hence, SC method has been widely used in natural image processing, such as image

* The National Natural Science Foundation of China (No. 60472111).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 220–226, 2008.
© Springer-Verlag Berlin Heidelberg 2008

feature extraction [6, 7] and image denoising [8, 9]. However, the standard SC algorithm described in [3] can not simultaneously guarantee the sparsity and independence of coefficient components. And the objective function itself can not balance between the sparseness constraint and the reconstructed precision. To remove or avoid the disadvantages mentioned above, in this paper, we propose a modified SC algorithm, which exploits the maximum Kurtosis as the maximizing sparse measure criterion, so the natural image structure captured by the Kurtosis not only is surely sparse, but also is surely independent. At the same time, a fixed variance term of coefficients is used to yield a fixed information capacity. This term can well balance the reconstructed error and sparsity. The experimental results also showed that utilizing our SC algorithm, the natural images' edge features can be extracted successfully. Further, applied these features, the original images can be reconstructed clearly.

2 The Modified Sparse Coding Algorithm

2.1 Modeling NNSC of Natural Images

Referring to the classical SC algorithm [3], and combining the minimum image reconstruction error with Kurtosis and fixed variance, we construct the following cost function of the minimization problem:

$$J(\mathrm{A},\mathrm{S}) = \frac{1}{2}\sum_{x,y}\left[\mathrm{X}(x,y) - \sum_i \mathrm{a}_i(x,y)\mathrm{s}_i\right]^2 - \lambda_1\sum_i\left|kurt(\mathrm{s}_i)\right| + \lambda_2\sum_i\left[\frac{\langle \mathrm{s}_i\rangle}{\sigma_t}\right]^2 . \tag{1}$$

where the symbol $\langle\cdot\rangle$ denotes the mean, $\mathrm{X}=(\mathrm{x}_1,\mathrm{x}_2,\ldots,\mathrm{x}_n)^T$ denotes the n-dimensional natural image data, $\mathrm{A}=(\mathrm{a}_1,\mathrm{a}_2,\ldots,\mathrm{a}_m)$ denotes the feature basis vectors, $\mathrm{S}=(\mathrm{s}_1,\mathrm{s}_2,\ldots,\mathrm{s}_m)^T$ denotes the m-dimensional sparse coefficients. In this paper, note that only the case of $m=n$ is considered, i.e., A is a square matrix. Parameters λ_1 and λ_2 are positive constant, σ_t^2 is the scale of coefficient variance. In Eqn. (1), the first term is the image reconstruction error and it ensures a good representation for a given image, and the second term is the sparseness measure based on the absolute value of Kurtosis, which is defined as:

$$\left|kurt(\mathrm{s}_i)\right| = \left|E\{\mathrm{s}_i^4\} - 3\left(E\{\mathrm{s}_i^2\}\right)^2\right| . \tag{2}$$

and maximizing $\left|kurt(s_i)\right|$ is equivalent to maximizing the sparseness of coefficient vectors; The last term can penalize the case in which the coefficient variance of the ith vector $\langle \mathrm{s}_i^2\rangle$ deviates from its target value σ_t^2. Without this term, the variance becomes so small that the sparseness constraint can only be satisfied, and the image reconstruction error would become large, which is not desirable either.

2.2 Learning Rules

Using the simple gradient descent algorithm to minimize the objective function, this differential equation of a_i is defined as follows:

$$\frac{\partial J(a_i, s_i)}{\partial a_i} = -\left[X - \sum_i a_i s_i\right] s_i^T \cdot \tag{3}$$

and further, the Eqn. (6) can be rewritten as follows:

$$a_i(k+1) = a_i(k) + \left[X - \sum_i a_i(k) s_i(k)\right] (s_i(k))^T \cdot \tag{4}$$

In a similar manner, the differential equation of s_i can be obtained as the following equation:

$$\frac{\partial J(a_i, s_i)}{\partial s_i} = -a_i^T \left[X - \sum_i a_i s_i\right] - \lambda_1 f_1(s_i) + \lambda_3 \frac{\ln\left(\langle s_i^2\rangle / \sigma_t^2\right)}{\langle s_i^2\rangle} \langle s_i\rangle \cdot \tag{5}$$

where $\lambda_3 = 4\lambda_2$ (a positive constant), $f_1(s_i) = \frac{\partial |kurt(s_i)|}{\partial s_i}$. According to Eqn.(2), the function $f_1(s_i)$ can be deduced as follows:

$$f_1(s_i) = \frac{\partial |kurt(s_i)|}{\partial s_i} = \beta\left[\langle s_i^3\rangle - 3\langle s_i^2\rangle \langle s_i\rangle\right] \cdot \tag{6}$$

where $\beta = sign(kurt(s_i))$, and for super-Gaussian signals, $\beta = 1$, and for sub-Gaussian signals, $\beta = -1$. Because of natural image data belonging to super-Gaussian, β is equal to 1. Thus, combined Eqn. (5) into Eqn. (6), then the updating rule of the coefficient variables can be obtained as follows:

$$s_i(k+1) = s_i(k) + (a_i(k))^T \left[X - \sum_i a_i(k) s_i(k)\right] + \lambda_4 \beta \left[\left\langle (s_i(k))^3 \right\rangle - 3\left\langle (s_i(k))^2 \right\rangle \langle s_i(k)\rangle\right] - \lambda_4 \langle s_i(k)\rangle . \tag{7}$$

where $\lambda_4 = 4\lambda_2 \left[\ln\left(\langle s_i^2\rangle / \sigma_t^2\right)\right] / \langle s_i^2\rangle$.

In performing loop, we update S and A in turn. First, holding A fixed, we update S, which is an inner loop. Then, holding S fixed, we update A, which is an outer loop. In order to improve the convergence speed, we use a determinative basis function, which is obtained by the independent component analysis (ICA), as the initialization feature basis function of our sparse coding algorithm instead of using a random initialization matrix. Note that the initialization feature bases obtained by ICA are not necessary to be converged to the least image reconstruction bias. Here, the loop times of the ICA method was assumed to be 100, and the learned bases were just used as the

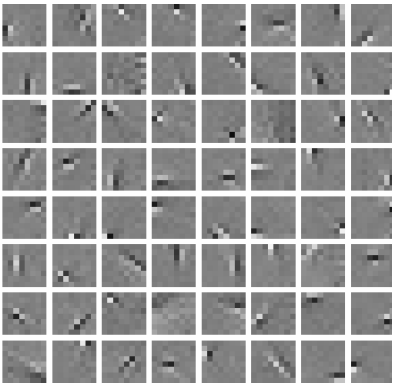


Fig. 1. Basis vectors obtained by applying our sparse coding algorithm to natural scenes. (a) Bases obtained by our algorithm; (b) Bases obtained by orthogonalized ICA.

initialization weight matrix. Using the above algorithm, the obtained results for 64 basis functions extracted from natural scenes are shown in Fig. 1 (a), where gray pixels denote zero weight, black pixels denote negative weight, and brighter pixels denote positive weights.

3 Experimental Results

3.1 Image Data Preprocessing

All test images used in our experiments are available on the following Internet web: http:// www.cns.nyu.edu/lcv/denoise. Firstly, selecting randomly 10 noise-free natural images with 512×512 pixels. Then, we sampled patches of 8×8 pixels 5000 times from each original image, and converted every patch into one column. Thus, each image was converted a matrix with 64×5000, and the size of input data set X consist of 10 test images was 64×50000 pixels. Consequently, each image patch is represented by a 8×8 dimensional vector. Further, the data set X was centered and whiten by the method of principal component analysis (PCA), and the preprocessed data set was denoted by $\hat{X}$. Then, using in turn the updating rules defined in Eqns. (4) and (7) of the feature basis matrix A and the sparse coefficient matrix A, the objective function given in Eqn. (1) was trained and the minimized results were obtained. The learned 64 feature basis vectors of natural scenes are shown in Fig. 1. as described in the subsection 2.2.

3.2 Image Reconstruction

Using the feature bases extracted, we can implement the task of restoring the original images. The test image was selected as the classical image shown in Fig. 2 (a), used widely in the image processing field, commonly known as Lena with 512×512 pixels. Assume that the number of image patches with 8×8 pixels sampled from Lena image

Fig. 2. The original image and reconstructed results obtained by our SC algorithm. (a) The original Lena image; (b) 5000 image patches; (c) 20000 image patches; (d) 50000 image patches.

was 5000, 20000 and 50000, respectively. Then, corresponding to different image patches, the image reconstruction results obtained by our SC algorithm were shown in Fig. 2 (b) to (d), in the same time, the reconstruction ones obtained by the standard SC method were also give, as shown in Fig. 3. Note that to find the accurate position of any image patch, we must remember the positions of all image patches sampled randomly. Because of sampling randomly, the same pixel might be founded in different image patches. Therefore, for the sample pixel, we averaged the sum of the values of all reconstructed pixels, and used the averaged pixel value as the approximation of the original pixel. Distinctly, from Fig.2 and Fig.3, it can be seen that the larger the number of image patches is, the clearer the reconstructed image is. When the number of image patches equals to 50000, it is difficult to tell it from the original images only with naked eyes. Moreover, compared Fig. 2 with Fig. 3, correspond to each reconstruction image with the same patches, respectively obtained by our SC algorithm and the standard SC algorithm, it is also not easy to tell them. Furthermore, this also testifies that our proposed SC algorithm can be indeed successfully applied on image feature extraction and image reconstruction.

In addition, the effectiveness of reconstruction can be further confirmed by the measure of the negative entropy H . And H is usually defined as:

$$H = -\sum_{i}^{255} p_i \log\left(p_i\right). \tag{8}$$

where p_i is the probability of the appearing gray pixels in an image. The calculated values of H of reconstructed images obtained by different methods were listed in Table 1. It is clear to see that the H value of any reconstructed image is significantly approximate to that of its original image, and the loss of image information is highly

(a) (b) (c)

Fig. 3. The reconstructed results obtained by standard SC algorithm. (a) 5000 image patches; (b) 20000 image patches; (c) 50000 image patches.

Table 1. Values of the negative entropy H obtained by different algorithms corresponding to different image patches

Algorithms	H of different Image patches			The original image
	5000	20000	50000	
Standard SC	7.4411	7.4432	7.4445	7.4451
Our SC	7.4425	7.4442	7.4448	

small. Therefore, our proposed SC algorithm can reserve the original image information as possible as. On the other hand, we also utilized the criterion of the objective fidelity to compare the reconstruction results implemented by methods of the standard SC and our SC. In general, the objective fidelity used usually in practice is the mean square signal to noise ratio (SNR) of the output image, denoted by SNR_{rms}. Let $I(x,y)$ denote the input image data set, and $\tilde{I}(x,y)$ denote the reconstruction image data, then the SNR_{rms} is defined as [10]:

$$\mathrm{SNR}_{rms} = \log_{10} \frac{\sum_{x=1}^{M}\sum_{y=1}^{N} \tilde{I}(x,y)^2}{\sum_{x=1}^{M}\sum_{y=1}^{N} \left[\tilde{I}(x,y) - I(x,y)\right]^2}. \tag{9}$$

where M and N denote the size of the image data, (x,y) is the pixel coordinate. Using the Eqn. (9), the calculated values of SNR_{rms} were listed in Table 2. Clearly, the SNR_{rms} value of each reconstructed image obtained by our method is litter greater than that obtained by the standard SC method. Therefore, the quality of reconstructed images of the former outperforms that of the latter.

Table 2. Values of the negative entropy H obtained by different algorithms corresponding to different image patches

Algorithms	SNR_{rms} of Image patches		
	5000	20000	50000
Standard SC	12.216	17.463	20.368
Our SC	13.352	19.862	22.853

4 Conclusions

In this paper, a novel natural image reconstruction method based on a modified sparse coding (SC) algorithm developed by us is proposed. This modified SC algorithm exploited the maximum Kurtosis as the maximizing sparse measure criterion, so the natural image structure captured by the Kurtosis not only is surely sparse, but also is surely independent. At the same time, a fixed variance term of coefficients is used to yield a fixed information capacity. Edge features of natural images can be extracted successfully by exploiting our SC algorithm. Simulation results showed that, compared with the standard SC algorithm, our proposed SC method is very effective in image feature extraction and image reconstruction based on values of SNR, which is used commonly in measuring image quality.

References

1. Cichocki, A.: Adaptive learning for principal component analysis with partial data. Vbernetics and Systems 2, 1014–1019 (1996)
2. Hyvärinen, A., Karhunen, J., Oja, E.: Independent component analysis. Wiley Interscience Publication, New York (2001)
3. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381, 607–609 (1996)
4. Lee, D.D., Seung, H.S.: Learning the parts of objects by non-negative matrix factorization. Nature 401(6755), 788–791 (1999)
5. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381, 607–609 (1996)
6. Bell, A., Sejnowski, T.J.: The 'Independent Components' of Natural Scenes Are Edge Filters. Vision Research 37, 3327–3338 (1997)
7. Hyvärinen, A., Oja, E., Hoyer, P., Horri, J.: Image Feature Extraction by Sparse Coding and Independent Component Analysis. In: Proc. Int. Conf. on Pattern Recognition (ICPR 1998), Brisbane, Australia, pp. 1268–1273 (1998)
8. Hyvärinen, A.: Sparse coding shrinkage: denoising of nongaussian data by maximum likelihood estimation. Neural Computation 11, 1739–1768 (1997)
9. Hyvärinen, A., Hoyer, P., Oja, E.: Image denoising by sparse code shrinkage. In: Haykin, S., Kosko, B. (eds.) Intelligent Signal Processing, pp. 554–568. IEEE Press, New York (2001)
10. Grgić, S., Grgić, M., Mrak, M.: Reliability of objective picture quality measures
11. Journal of Electrical Engineering. 55, 3–10 (2004)

Application of Optimization Technique for GPS Navigation Kalman Filter Adaptation

Dah-Jing Jwo and Shun-Chieh Chang

Department of Communications, Navigation and Control Engineering
National Taiwan Ocean University, Keelung 20224, Taiwan
djjwo@mail.ntou.edu.tw

Abstract. The position-velocity (PV) process model can be applied to the GPS Kalman filter adequately when navigating a vehicle with constant speed. However, when an abrupt acceleration motion occurs, the filtering solution becomes very poor or even diverges. To avoid the limitation of the Kalman filter, the particle swarm optimization can be incorporated into the filtering mechanism as dynamic model corrector. The PSO can be utilized as the noise-adaptive mechanism to tune the covariance matrix of process noise and overcome the deficiency of Kalman filter. In this paper, PSO-aided Kalman filter approach is employed for tuning the covariance of the GPS Kalman filter so as to reduce the estimation error during substantial maneuvering. Performance evaluation for the PSO-aided Kalman filter as compared to the conventional Kalman filter is provided.

Keywords: GPS, Particle swarm optimization, Adaptive Kalman filter.

1 Introduction

The Global positioning systems (GPS) [1,2] are capable of providing accurate position information. While employed in the GPS receiver as the navigational state estimator, the extended Kalman filter (EKF) [1-3] has been one of the promising approaches. To obtain good estimation solutions using the EKF approach, the designers are required to have good knowledge on both dynamic process (plant dynamics, using an estimated internal model of the dynamics of the system) and measurement models, in addition to the assumption that both the process and measurement are corrupted by zero-mean white noises. The divergence due to modeling errors is critical in Kalman filter designs. If the input data does not reflect the real model, the KF estimates may not be reliable. In various circumstances where there are uncertainties in the system model and noise description, and the assumptions on the statistics of disturbances are violated since in a number of practical situations, the availability of a precisely known model is unrealistic due to the fact that in the modeling step, some phenomena are disregarded and a way to take them into account is to consider a nominal model affected by uncertainty.

In navigation filter designs, there exist the model uncertainties which cannot be expressed by the linear state-space model. The linear model includes modeling errors since the actual vehicle motions are non-linear process. The system model, system

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 227–234, 2008.
© Springer-Verlag Berlin Heidelberg 2008

initial conditions, and noise characteristics have to be specified *a priori*. It is very often the case that little *a priori* knowledge is available concerning the maneuver. To fulfill the requirement of achieving the filter optimality or to preventing divergence problem of Kalman filter, the so-called adaptive Kalman filter (AKF) approach [4-6] has been one of the promising strategies and has been widely explored for dynamically adjusting the parameters of the supposedly optimum filter based on the estimates of the unknown parameters for on-line estimation of motion as well as the signal and noise statistics available data. The innovation-based adaptive estimation (IAE) approach has been very popular in AKF. The AKF can be used for improving the tracking capability in high dynamic maneuvering.

Particle swarm optimization (PSO) [7,8] is a population based stochastic searching technique developed by Kennedy and Eberhart [7]. It is a relatively recent heuristic search method whose mechanics are inspired by swarming or collaborative behavior of biological populations. Among various evolutionary optimizer techniques, Genetic Algorithms (GA) and PSO have attracted considerable attention. The method proposed in this paper makes use of the particle swarm optimization (PSO) techniques, which is employed into the GPS navigation filter for real-time identification of noise covariance matrices to prevent divergence of the Kalman filter.

2 The Adaptive Kalman Filter

The discrete-time Kalman filter algorithm is summarized as follow:

Prediction steps/time update equations:

$$\hat{\mathbf{x}}_{k+1}^{-} = \mathbf{\Phi}_k \hat{\mathbf{x}}_k \,. \tag{1}$$

$$\mathbf{P}_{k+1}^{-} = \mathbf{\Phi}_k \mathbf{P}_k \mathbf{\Phi}_k^{\mathrm{T}} + \mathbf{\Gamma}_k \mathbf{Q}_k \mathbf{\Gamma}_k^{T} \,. \tag{2}$$

Correction steps/measurement update equations:

$$\mathbf{K}_k = \mathbf{P}_k^{-} \mathbf{H}_k^{\mathrm{T}} [\mathbf{H}_k \mathbf{P}_k^{-} \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k]^{-1} \,. \tag{3}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^{-} + \mathbf{K}_k [\mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^{-}] \,. \tag{4}$$

$$\mathbf{P}_k = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_k^{-} \,. \tag{5}$$

The innovation sequences have been utilized by the correlation and covariance-matching techniques to estimate the noise covariances. The basic idea behind the covariance-matching approach is to make the actual value of the covariance of the residual consistent with its theoretical value. The implementation of IAE based AKF to navigation designs has been widely explored [4-6]. From the incoming measurement $\mathbf{z}_k$ and the optimal prediction $\hat{\mathbf{x}}_k^{-}$ obtained in the previous step, the innovations sequence is defined as

$$\boldsymbol{\upsilon}_k = \mathbf{z}_k - \hat{\mathbf{z}}_k^{-} = \mathbf{H}_k (\mathbf{x}_k - \hat{\mathbf{x}}_k^{-}) + \mathbf{v}_k \,. \tag{6}$$

The innovation reflects the discrepancy between the predicted measurement $\mathbf{H}_k\hat{\mathbf{x}}_k^-$ and the actual measurement $\mathbf{z}_k$. It represents the additional information available to the filter as a consequence of the new observation $\mathbf{z}_k$. The innovations sequence $\boldsymbol{\upsilon}_k$ is a zero-mean Gaussian white noise sequence. An innovation of zero means that the two are in complete agreement. The mean of the corresponding error of an unbiased estimator is zero. By taking variances on both sides, we have the theoretical covariance, the covariance matrix of the innovation sequence is given by

$$\mathbf{C}_{\upsilon_k} = E[\boldsymbol{\upsilon}_k\boldsymbol{\upsilon}_k^{\mathrm{T}}] = \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k \,. \tag{7a}$$

which can be written as

$$\mathbf{C}_{\upsilon_k} = \mathbf{H}_k(\boldsymbol{\Phi}_k\mathbf{P}_k\boldsymbol{\Phi}_k^T + \boldsymbol{\Gamma}_k\mathbf{Q}_k\boldsymbol{\Gamma}_k^T)\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k \,. \tag{7b}$$

Defining $\hat{\mathbf{C}}_{\upsilon_k}$ as the statistical sample variance estimate of $\mathbf{C}_{\upsilon_k}$, matrix $\hat{\mathbf{C}}_{\upsilon_k}$ can be computed through averaging inside a moving estimation window of size N

$$\hat{\mathbf{C}}_{\upsilon_k} = \frac{1}{N}\sum_{j=j_0}^{k}\boldsymbol{\upsilon}_j\boldsymbol{\upsilon}_j^{\mathrm{T}} \,. \tag{8}$$

where N is the number of samples (usually referred to the window size); $j_0 = k - N + 1$ is the first sample inside the estimation window. The window size N is chosen empirically (a good size for the moving window may vary from 10 to 30) to give some statistical smoothing. More detailed discussion can be referred to Brown & Hwang [1] and Gelb [3]. An adaptive Kalman filter can be utilized as the noise-adaptive filter to estimate the noise covariance matrices and overcome the deficiency of Kalman filter. The benefit of the adaptive algorithm is that it keeps the covariance consistent with the real performance. The innovation sequences have been utilized by the correlation and covariance-matching techniques to estimate the noise covariances. The basic idea behind the covariance-matching approach is to make the actual value of the covariance of the residual consistent with its theoretical value. For more detailed information derivation for these equations, see Mohamed & Schwarz [5].

3 The Proposed PSO-Aided Adaptive Kalman Filter

To account for the greater uncertainty, Kalman filtering with motion detection can be incorporated into GPS navigation filter for timely updated the covariance, through the way:

$$\mathbf{Q}_k \rightarrow \boldsymbol{\lambda}_k\mathbf{Q}_k \,. \tag{9}$$

Substituting Equation (9) into Equation (4) leads to

$$\mathbf{P}_{k+1}^- = \boldsymbol{\Phi}_k\mathbf{P}_k\boldsymbol{\Phi}_k^{\mathrm{T}} + \boldsymbol{\lambda}_k\boldsymbol{\Gamma}_k\mathbf{Q}_k\boldsymbol{\Gamma}_k^T \,. \tag{10}$$

Equation (10) is utilized in new Kalman filter formulation. In case that $\boldsymbol{\lambda}_k = \mathbf{I}$, it deteriorates to the standard Kalman filter. Fig. 1 shows the basic philosophy for the PSO-aided Kalman filter algorithm.

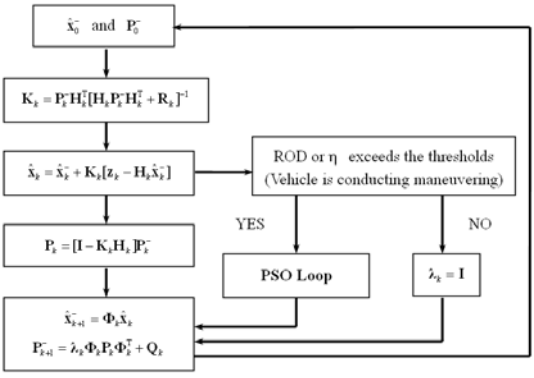


Fig. 1. Flow chart for the PSO-aided Kalman filter algorithm

3.1 Particle Swarm Optimization (PSO)

Particle swarm optimization (PSO) is a population based stochastic searching technique developed by Kennedy and Eberhart [7]. It is a relatively recent heuristic search method whose mechanics are inspired by swarming or collaborative behavior of biological populations. Among various evolutionary optimizer techniques, Genetic Algorithms (GA) and PSO have attracted considerable attention [8]. The PSO is a robust stochastic evolutionary computation technique based on the movement and intelligence of swarms looking for the most fertile feeding location. Unlike the drawback of expensive computational cost of GA, PSO has better convergence speed.

A *swarm* consists of a set of particles moving around the search space, each representing a potential solution (fitness). Each particle has a position vector (x_i), a velocity vector (v_i), the position at which the best fitness ($Pbest_i$) encountered by the particle, and the index of the best particle (*Gbest*) in the swarm. The position of each particle is updated every generation. This is done by adding the velocity to the position vector.

$$v_i = v_i + C_1 \times rand() \times (Pbest_i - x_i) + C_2 \times rand() \times (Gbest - x_i) . \tag{11}$$

The positions are based on their movement over a discrete time interval (Δt) as follows, with Δt usually set to 1.

$$x_i = x_i + v_i \cdot \Delta t . \tag{12}$$

The parameters C_1 and C_2 are set to positive constant values, which are normally taken as 2 whereas rand() represent uniformly distributed random values, uniformly distributed in [0, 1] and *w* is called as inertia weight, the inertia weight is employed to control the impact of the previous history of velocities on the current one.

3.2 PSO-Aided Kalman Filter

The innovation information at the recent epochs is collected for timely reflect the change in vehicle dynamic. Defining the rate of divergence (ROD) as the trace of innovation covariance matrix, we have:

$$ROD = \frac{tr(\hat{\mathbf{C}}_{\upsilon_k})}{tr(\mathbf{C}_{\upsilon_k})}. \tag{13}$$

This parameter can be utilized for divergence/outliers detection or parameter adaptation in adaptive filtering. In addition, the other parameter for detecting the high dynamic motion is given by

$$\eta = \frac{1}{N} \sum_{j=j0}^{k} (\Delta \hat{\mathbf{x}}_k). \tag{14}$$

where

$$\Delta \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k - \hat{\mathbf{x}}_k^- = \mathbf{K}_k \upsilon_k. \tag{15}$$

Here η is employed to assist ROD to detect the divergence (high dynamic motion) in PSO-aided KF. The two parameters are employed for identifying the degree of dynamical change of vehicle motion. In the PSO-aided KF, the ROD parameter is employed again as the fitness function (FIT) in the PSO loop. The optimization problem is performed using PSO by iteratively tune the FIT parameter

$$FIT = \frac{tr(\hat{\mathbf{C}}_{\upsilon_k})}{tr(\mathbf{H}_k(\boldsymbol{\Phi}_k \mathbf{P}_k \boldsymbol{\Phi}_k^T + \boldsymbol{\lambda}_k \boldsymbol{\Gamma}_k \mathbf{Q}_k \boldsymbol{\Gamma}_k^T)\mathbf{H}_k^T + \mathbf{R}_k)}. \tag{16}$$

so as to keeps the predicted covariance consistent with the actual one from sampled sequence. When the required FIT is obtained, iteration is terminated and the $\boldsymbol{\lambda}_k$ is determined. The flow chart of the PSO algorithm is given by Fig. 2.

4 Application to GPS Navigation Processing

Simulation experiments have been carried out to evaluate the performance of the proposed approach in comparison with the conventional methods. Assume that the differential GPS mode is used and most of the errors can be corrected, but the multipath and receiver thermal noise cannot be eliminated. The dynamic process of the GPS receiver in medium dynamic environment can be represented by the PV (Position-Velocity) model [1,2]. In such case, the state to be estimated is a 8×1 vector. including three position components, three velocity components, and the receiver clock offset and drift errors, respectively. The states are implemented in the WGS-84 coordinates.

The trajectory and the corresponding east and north components of velocity for the simulated vehicle is shown in Fig. 3. Experiment was conducted on a simulated vehicle trajectory originating from the (0, 0) m location. The trajectory of the vehicle can be approximately divided into two categories according to the dynamic characteristics. The vehicle was simulated to conduct constant-velocity straight-line during the three time intervals, 0-250, 751-1250 and 1751-2000s, all at a speed of 10π m/s. Furthermore, it conducted circular motion with radius 2500 meters during 251-750 (counterclockwise), and 1251-1750s (clockwise) where high dynamic maneuvering is involved.

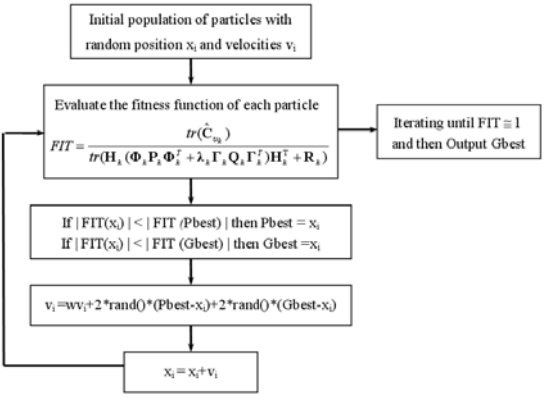


Fig. 2. Flow chart of the PSO algorithm

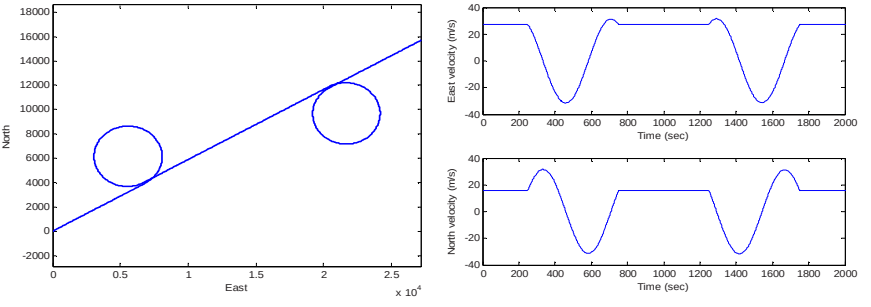


Fig. 3. Trajectory (left) and the corresponding east and north components of the vehicle velocity (right) for the simulated vehicle

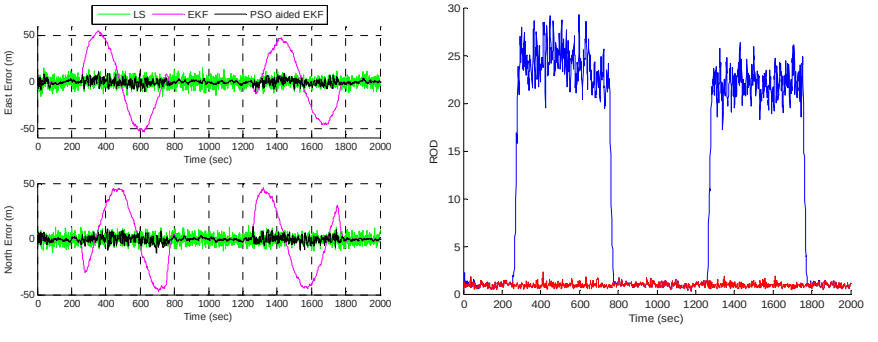


Fig. 4. GPS navigation solution without and with adaptation (left); comparison of ROD trajectories with and without PSO aiding (right)

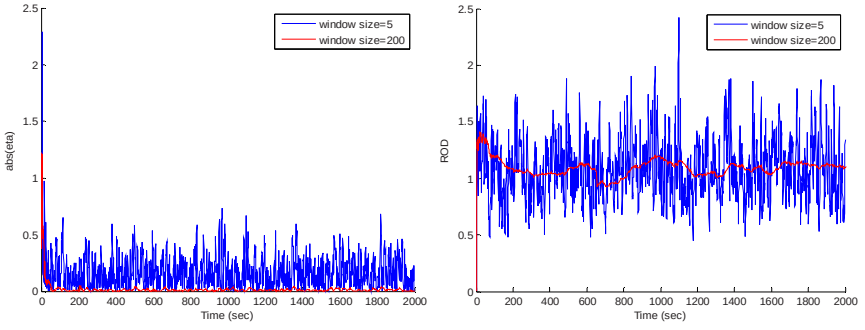


Fig. 5. η (left) and ROD (right) trajectories for window size 5 and 200

Fig. 4 provides GPS navigation solutions without and with adaptation as compared to that by least-squares (LS) approach, where the parameters used are: $ROD > 1.7$, $|\eta| > 0.5$, window size 5. **C**omparison of ROD trajectories with and without PSO aiding is also provided. Substantial improvement in tracking capability can be obtained. Some principles for the setting of thresholds and window size are discussed. Taking the limit for Equation (14) leads to

$$\lim_{N\to\infty}\left|\frac{1}{N}\sum_{j=j_0}^{k}(\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_k^-)\right| \approx 0\,. \tag{17}$$

The two parameters has the relation

$$N \propto \frac{1}{|\eta|}\,. \tag{18}$$

Therefore, when the window size is set to a larger value, the threshold of η can be set to a smaller one. When the window size is increased, $\eta\,(\geq 0)$ is decreased toward 0,

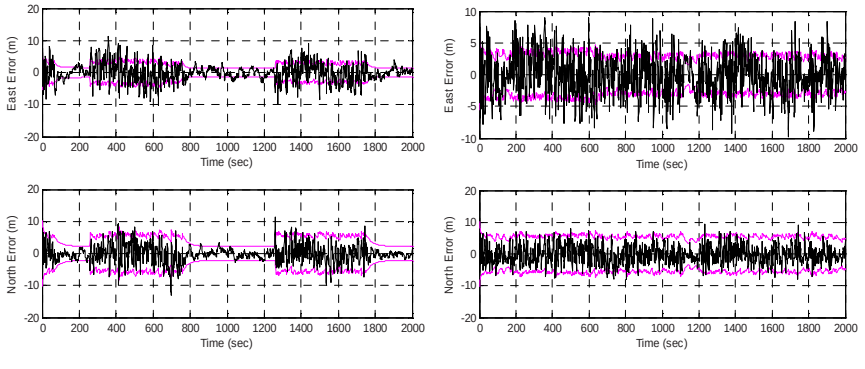


Fig. 6. East and north components of navigation errors and the 1-σ bound using the proposed method: $ROD > 1.7$, $|\eta| > 0.5$) (left); $ROD > 1$, $|\eta| > 0.1$) (right)

while ROD is increased toward 1, meaning that $tr(\hat{\mathbf{C}}_{\upsilon_k})$ is approaching to $tr(\mathbf{C}_{\upsilon_k})$. Fig. 5 shows η and ROD trajectories for window size 5 and 200, respectively. The value of window size will influence the adaptation performance, and accordingly the estimation accuracy. Fig. 6 shows the east and north components of navigation errors and the 1-σ bound for $ROD > 1.7$, $|\eta| > 0.5$, and for $ROD > 1$, $|\eta| > 0.1$, respectively.

5 Conclusions

This paper has presented a PSO aided KF method for GPS navigation processing, where the conventional KF approach is coupled by the adaptive tuning system constructed by the PSO. The PSO provides the process noise covariance scaling factor for timely detecting the dynamical and environmental changes and implementing the online parameter tuning by monitoring the innovation information so as to maintain good tracking capability and estimation accuracy. The proposed method has the merits of good numerical stability since the matrices in the KF loop are able to remain positive definitive. Simulation experiments for GPS navigation have been provided to illustrate the accessibility. In addition, behavior of some innovation related parameters have been discussed, which are useful for providing the useful information in designing the adaptive Kalman filter and for achieving the system integrity. The navigation accuracy based on the proposed method has been compared to the conventional EKF method and has demonstrated substantial improvement in both navigational accuracy and tracking capability.

Acknowledgments. This work has been supported in part by the National Science Council of the Republic of China under grant no. NSC 96-2221-E-019-007.

References

1. Brown, R., Hwang, P.: Introduction to Random Signals and Applied Kalman Filtering. John Wiley & Sons, New York (1997)
2. Farrell, J.: The Global Positioning System and Inertial Navigation. McGraw-Hill professional, New York (1998)
3. Gelb, A.: Applied Optimal Estimation. MIT Press, MA (1974)
4. Mehra, R.K.: Approaches to adaptive filtering. IEEE Trans. Automat. Contr. AC-17, 693–698 (1972)
5. Mohamed, A.H., Schwarz, K.P.: Adaptive Kalman Filtering for INS/GPS. Journal of Geodesy 73, 193–203 (1999)
6. Hide, C., Moore, T., Smith, M.: Adaptive Kalman Filtering for Low Cost INS/GPS. Journal of Navigation 56, 143–152 (2003)
7. Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. In: Proc. IEEE International conf. Neural Network, pp. 1942–1945 (1995)
8. Eberhart, R.C., Shi, Y.: Comparison Between Genetic Algorithms and Particle Swarm Optimization. In: 1998 Annual Conf. on Evolutionary Programming (1998)

Optimal Design of Passive Power Filters Based on Multi-objective Cultural Algorithms

Yi-nan Guo, Jian Cheng, Yong Lin, and Xingdong Jiang

School of Information and Electronic Engineering, China University of Mining and Technology, Xuzhou, 221008 Jiangsu, China
nanfly@126.com

Abstract. Design of passive power filters shall meet the demand of harmonics suppression effect and economic target. However, existing optimization methods for this problem only take technology target into account or do not utilize knowledge enough, which limits the speed of convergence and the performance of solutions. To solve the problem, two objectives including minimum total harmonics distortion of current and minimum cost for equipments are constructed. In order to achieve the optimal solution effectively, a novel multi-objective optimization method, which adopts dual evolution structure in cultural algorithms, is adopted. Implicit knowledge describing the dominant space are extracted and utilized to induce the direction of evolution. Taken three-phase full wave controlled rectifier as harmonic source, simulation results show that filter designed by the proposed algorithm have better harmonics suppression effect and lower investment for equipments than filter designed by existing methods.

Keywords: Passive power filter, cultural algorithm, multi-objective, harmonics.

1 Introduction

With the wide use of nonlinear power electronics equipments, harmonic pollution to power system caused by these loads is increasingly severe. Harmonics may do harm to the quality of power system, such as overheat of electric apparatus, false action of relay protection and so on. Meanwhile, more and more loads demand better quality of power system. In order to suppress harmonics effectively, hybrid power filters including passive power filters (PPFs) and active power filters (APFs) are used. PPFs are the necessary part used to realize reactive power compensation. APFs are used to improve harmonics suppression effect. So design of PPF is an important issue which influences the total performance of hybrid power filters.

The goal of design is to achieve a group of PPF's parameters which have better harmonics suppression effect, lower cost and reasonable capacity of reactive power compensation. Traditional experience-based design method has difficulty achieving the most rational parameters because it only considers harmonics suppression effect of PPF[1]. Aiming at the problem, many multi-objective optimization methods considering economic and technology targets are introduced. First, weighted multi-objective method is adopted to simplify the problem into single objective.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 235–242, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Based on the simplified objective, some existing evolutionary algorithms can be used directly, such as genetic algorithm[2], simulated annealing[3]. Second, ε-constraint method is used to consider harmonics suppression effect and cost simultaneously [4]-[6]. But it is hard to decide ε suitably. Third, pareto-based methods are introduced, such as NSGA-II[7]. Although these methods can find the satisfied solution, the performance of algorithms are limited because the knowledge embedded in evolution process is not fully used.

Aiming at the problem, taken best harmonics suppression effect and lowest cost as optimization objectives, a novel optimal design method adopting multi-objective cultural algorithm (MCA) is proposed. In the method, implicit knowledge describing dominant space are extracted and utilized to induce the direction of evolution.

2 Description of the Problem

The principle of PPF is that the *k*-th harmonic current will flow into the filter and be eliminated from power system when the frequency of the *k*-th filtered harmonic equals to the resonant frequency of PPF. According to the order of filtered harmonic, PPFs normally consist of single-tuned filter and high pass filter, which are composed of resistor(R), inductor (L) and capacitor(C).

In single-tuned filter, the reactance of inductor is equal to that of capacitor at resonant frequency. Suppose w_k is resonant frequency. $Q \in [30,60]$ is the quality factor. The relationship among R_k, L_k, C_k at resonant frequency is expressed in (1).

$$L_k = \frac{1}{w_k^2 C_k} \qquad R_k = \frac{w_k L_k}{Q} \tag{1}$$

Suppose w_h is cut-off frequency. m=0.5 is damped-time-constant-ratio.Based on the condition of resonant, the relationship among R_h, L_h, C_h in high pass filters are

$$R_h = \frac{1}{w_h C_h} \qquad L_h = \frac{R_h m}{w_h} \tag{2}$$

Considering harmonics suppression effect and economic target of PPFs, the objective functions of the problem include total harmonic distortion of current (or voltage) and cost.

1) Harmonics suppression effect

A good passive power filter shall have lower total harmonics distortion of voltage THD_u or current THD_i which reflect harmonics suppression effect. That is

$$\min \ THD_u = \sqrt{\sum_{k=1}^{\infty} \left(\frac{U_k}{U_1} \right)^2} \times 100\ \% \tag{3}$$

$$\min \ THD_i = \sqrt{\sum_{k=1}^{\infty} \left(\frac{I_k}{I_1} \right)^2} \times 100\ \% \tag{4}$$

where U_1 and I_1 denote the fundamental voltage and current.U_k, I_k denote the *k*-th order harmonic voltage and current. Let *l* is the order of the single-tuned filters.

$$U_k = \frac{I_k}{Y_{sk} + \sum_l Y_{kl} + Y_{hk}} \qquad I_k = \frac{1/\sum_l Y_{kl} + Y_{hk}}{Z_{sk} + \sum_l Z_{kl} + Z_{hk}} \tag{5}$$

$$Z_{sk} = R_s + jw_k L_s, \quad Z_{kl} = R_l + j\left(w_k L_l - \frac{1}{w_k C_l}\right), \quad Z_{hk} = R_h \,//\, jw_k L_h + \frac{1}{jw_k C_h} \tag{6}$$

where Z_{sk} is the equivalent k-th harmonic impedance of power system. Z_{kl} and Z_{hk} are the equivalent k-th harmonic impedance of each single-tuned filter and high pass filter. Y_{sk} ,Y_{kl} and Y_{hk} are the adpedance of equivalent harmonic impedance.

2) cost

Cost directly influences the economic benefit of any enterprises. For PPFs, early investment for equipments, which is decided by the unit prices of C,L,R, is more important because the electric power capacitor is expensive. Let q_1, q_2, q_3 are the unit price of R_j, L_j, C_j. The cost of PPF is expressed as follows.

$$\min Co = \sum_j (q_1 R_j + q_2 L_j + q_3 C_j) \tag{7}$$

3) reactive power compensation

In power system with PPFs, high reactive power compensation is expected. But reactive power can't be overcompensation as power factor is close to 1. Suppose the capacity of reactive power compensation is $Q_j = \omega C_j U_j^2$. Above need is expressed by

$$\max \sum_j Q_j \quad s.t. Q_{\min} \le \sum_j Q_j \le Q_{\max} \tag{8}$$

It is obvious that optimal design of PPF is a dual-objective optimization problem with one constraint. Good harmonics suppression effect needs good equipments which will increase cost. So above objectives are non-dominant each other.

3 Optimal Design of PPF Based on Multi-objective Cultural Algorithms

The goal of design is to find a group of PPF's parameters which have lowest cost and best harmonics suppression effect. According to the relationship among R,L,C, it is obvious that as long as the values of Q, m and capacitors in filters are decided, the values of other parameters can be achieved by computation. Moreover, electric power capacitor is most expensive. So capacitors in single-tuned filters or high pass filter are chosen as variables. In order to reduce the cumulative error of encoding, real number is adopted to describe each variable in an individual.

It has been proved that implicit knowledge extracted from evolution process can be used to induce evolution process so as to improve the performance of algorithms. Based on above theory, multi-objective cultural algorithm with dual evolution structure is adopted in the problem, as shown in fig.1. In population space, evolution

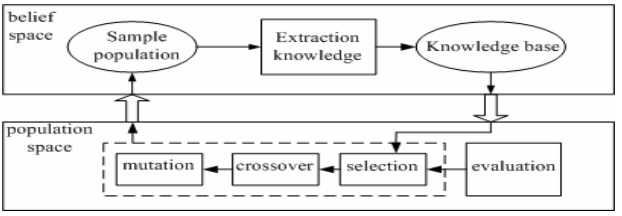


Fig. 1. The structure of optimal design method based on multi-objective cultural algorithm

operators, including evaluation, selection, crossover and mutation, are realized. Here,NAGA-II is adopted in population space. The individuals dominating others or non-comparing with others are saved as samples in sample population. In belief space, implicit knowledge is extracted according to the fitness of these samples and utilized in the selection process.

1) initialization : Logistic chaotic sequence $x_i(t+1)=\mu x_i(t)$ show the ergodicity in [0,1] when μ=4. So it is introduced into initialization in order to ensure the diversity of initial population. Firstly, random numbers $\tilde{c}_{1j}(0)\in[0,1], j=1,2,\cdots,n$ are generated where *n* is the number of filters.Secondly, each random number is used as the initial value of a chaotic sequence. Different Logistic chaotic sequences are generated based on different random number. And the length of each sequence equals to population size |*P*|. Third, each number in this |*P*|×*n* matrix is transformed into the real number satisfied the capacity limit of electric power capacitor,shown as $c_{ij}(0)\in[\underline{C_j},\overline{C_j}], i=1,2,\cdots,m\cdot\underline{C_j}$, $\overline{C_j}$ denotes lower and upper limits of capacitors.

2) extraction and utilization of implicit knowledge: In multi-objective cultural algorithm, implicit knowledge describes the distribution of non-dominated individuals in objective space. The objective space is formed according to the bound of each objective function $\Omega:\{(f_1(X),f_2(X))|\ f_1(X)\in[\underline{f_1},\overline{f_1}], f_2(X)\in[\underline{f_2},\overline{f_2}]\}$. Then it is divided into several uniform cells $\Omega c_k\subset\Omega, k=1,2,\cdots,N_\Omega\times N_\Omega$ along each objective function [8]. If the non-compared or dominating individuals locating in a cell are more, the degree of this cell is larger. Suppose |Ps| is the sampling population size. The degree of *k*-th cell is $D_{\Omega c_k}=N_{c_k}, s.t.\sum_{k=1}^{N_\Omega\times N_\Omega}N_{c_k}=|Ps|$. It is obvious that $D_{\Omega c_k}$ reflects the crowd degree of individuals in *k*-th cell.

The uniform distribution of optimal pareto front is an important performance in multi-objective optimization problem. Existing methods normally adopt crowd operator or niche sharing fitness to maintain the uniform distribution of pareto front. However, the niche radius is difficult to given. Aiming at the problem, implicit knowledge is used to replace crowd operator. Combing with tournament selection, the rules for selecting individuals are shown as follows.

Rule1: An individual dominating others wins.

Rule2: The non-dominated individual located in the cell with lower crowd degree wins.

4 Simulation Results and Their Analysis

4.1 Analysis of the System with Harmonic Source

In the paper, three-phase full wave controlled rectifier is adopted as the harmonic source, shown in fig.2. Because phases controlled rectification bridge is a normal harmonic current source, its harmonic distortion of voltage is small. The wave and spectrum character of U_1 and I_1 in above system without PPF are shown in fig.2.The harmonic current without PPF include 5th, 7th, 11th, 13th, 17th, 19th order harmonics and so on. And the harmonics, which order are larger than 20, are ignored.

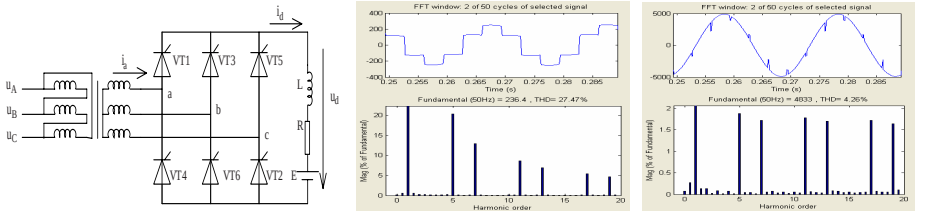


Fig. 2. Three-phase full wave controlled rectifier and the wave and spectrum of I_1 and U_1 without PPF

It is obvious that in above power system without PPF, the contents of 5th, 7th, 11th, 13th order harmonic are large. Total harmonics distortion of current and voltage is 27.47% and 4.26%. So PPF for above harmonic source consists of four single-tuned filters for 5th, 7th, 11th, 13th order harmonic and a high pass filter with cut-off frequency near 17th order harmonic. Suppose the equivalent 5th, 7th, 11th, 13th harmonic impedance of single-tuned filters and high pass filter are

$$Z_5 = R_5 + \frac{j}{C_5}\left(\frac{n}{25w} - \frac{1}{nw}\right)' Z_7 = R_7 + \frac{j}{C_7}\left(\frac{n}{49w} - \frac{1}{nw}\right) \tag{9}$$

$$Z_{11} = R_{11} + \frac{j}{C_{11}}\left(\frac{n}{121w} - \frac{1}{nw}\right)' Z_{13} = R_{13} + \frac{j}{C_{13}}\left(\frac{n}{169w} - \frac{1}{nw}\right)' Z_h = \frac{jnR_h}{34 + jn} - \frac{j17R_h}{n} \tag{10}$$

The total impedance of system is introduced from above harmonic impedance, shown as $Z = \frac{Z_5 \bullet Z_7 \bullet Z_{11} \bullet Z_{13} \bullet Z_h}{Z_5 \bullet Z_7 \bullet Z_{11} \bullet Z_{13} + Z_5 \bullet Z_7 \bullet Z_{11} \bullet Z_h + Z_5 \bullet Z_7 \bullet Z_{13} \bullet Z_h + Z_5 \bullet Z_{11} \bullet Z_{13} \bullet Z_h + Z_7 \bullet Z_{11} \bullet Z_{13} \bullet Z_h}$. Therefore, an optimization objective about total harmonics distortion of current is given by (11).

$$\min THD_i = \sqrt{\sum_{j=5,7,11,13,h}\left(\frac{I_j}{I_1}\right)^2} \times 100\ \% = \sqrt{\sum_{j=5,7,11,13,h}\left(\frac{Z}{kZ_j}\right)^2} \times 100\ \% \tag{11}$$

According to resonate theory, resistor and inductor in formula (7) can be transformed into capacitor. So simplified function about cost is

$$Co = \sum_{j=5,7,11,13} \left(\frac{q_1}{w_i C_i Q} + \frac{q_2}{w_i^2 C_i} + q_3 C_i\right) + \left(\frac{q_1}{w_h C_h} + \frac{q_2 m}{w_h^2 C_h} + q_3 C_h\right) \tag{12}$$

4.2 Simulation Tests

In order to analyze harmonics suppression effect and cost of PPF optimized by the proposed algorithm, the main parameters used in the algorithm are shown in Tab.1.

Table 1. The parameters in the algorithm

Termination iteration	Probability of crossover	Probability of mutation	$\|P\|$	Q	m
100	0.9	0.1	50	50	0.5
unit price of R	unit price of L	unit price of C	Bound of C_h		
¥ 88/Ω	¥ 320/mH	¥ 4800/μF	[25,35]		
Bound of C_5	Bound of C_7	Bound of C_{11}	Bound of C_{13}		
[45,65]	[25,35]	[25,35]	[25,35]		

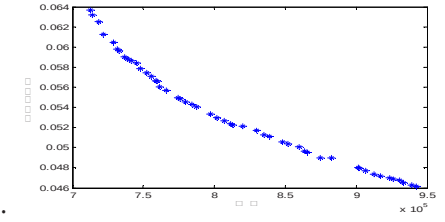


Fig. 3. The relationship between two objectives

Based on above parameters, the pareto front, which reflects the relationship between total harmonics distortion of current and cost, are shown in fig.3. It is obvious that these objectives are non-dominant each other which accord with the trait of min-min pareto front. That is, if the decisionmakers think economic target is more important, expense for equipments decreases, whereas total harmonics distortion of current becomes worse.

In order to validate the rationality of MCA, THD_i, THD_u and cost obtained by MCA are compared with Ref[9] and traditional experience-based design method(EBM) aiming at the power system with above load. Considering the different importance of two objectives, three groups of PPF's optimal parameters and corresponding objectives values are achieved. Here, weight coefficients in the method proposed in Ref[9] are 1)ω_1=1, ω_2=0, 2)ω_1=0.5, ω_2=0.5, 3)ω_1=0, ω_2=1 respectively. Corresponding THD_i, THD_u and cost are listed in Table.2. Comparison of the results show that the performances of PPF designed by MCA are better than them of Ref[9] and experience-based solution. Traditional experience-based design

Table 2. Comparison of the performance among MCA, Ref[9]and experience-based method

	method	order	R (Ω)	L (mH)	C (μF)	THD_i (%)	THD_u (%)	cost (10^5¥)
1	MCA	5	0.19598	6.2415	65	**5.43**	**1.34**	**9.395**
		7	0.26719	6.078	34.055			
		11	0.16564	2.3977	34.959			
		13	0.14027	1.7182	34.929			
		high pass	7.3462	0.68811	25.501			
	Ref[9]	-	-	-	-	6.17	1.56	9.571
	EBM	-	-	-	-	8.28	2.04	9.892
2	MCA	5	0.24696	7.865	51.582	**6.57**	**1.45**	**8.116**
		7	0.33805	7.6899	26.917			
		11	0.2006	2.8961	28.943			
		13	0.14	1.7149	34.996			
		high pass	7.4625	0.699	25.104			
	Ref[9]	-	-	-	-	7.46	1.83	8.723
	EBM	-	-	-	-	-	-	-
3	MCA	5	0.28293	9.0107	45.024	**8.36**	**1.9**	**7.125**
		7	0.36098	8.2117	25.207			
		11	0.23162	3.3529	25			
		13	0.19598	2.4006	25			
		high pass	7.0754	0.66274	26.477			
	Ref[9]	-	-	-	-	9.69	2.38	7.806
	EBM	-	-	-	-	-	-	-

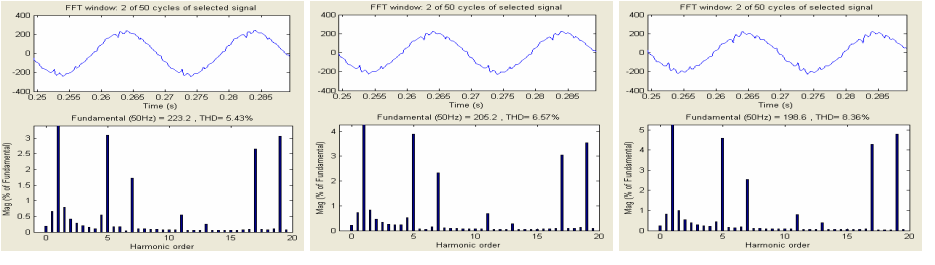


Fig. 4. The wave and spectrum of I_1 with PPT designed by MCA in Tabel.2

method for PPF only considers harmonics suppression effect. So it is difficult to find the optimal solutions. In Ref[9], though two objectives are both considered, they are transformed into single objective which can only provide discrete solutions for the decisionmakers. The wave and spectrum of I_1 with PPT designed by MCA under different importance of two objectives are shown in fig.4. Though harmonics are suppressed effectively, total harmonics distortion of current become larger gradually along with the increasing importance for cost. This accords with the min-min pareto front shown in fig.3.

4 Conclusions

PPF is the important device for harmonic suppression. Design of PPF is to obtain a group of parameters including resistor, inductor and capacitor which meet the demand of technology and economic targets. However, existing method only considers harmonic suppression effect or do not utilize knowledge enough, which limits the speed of convergence and the performance of solutions. Aiming at the problem, two objectives including minimum total harmonics distortion of current and minimum cost of PPF are considered. Multi-objective cultural algorithm with dual evolution structure using capacitors as variables is adopted. Taken three-phase full wave controlled rectifier as harmonic source, simulation results show that PPF designed by the proposed algorithm has better harmonics suppression effect and lower investment for equipments than filter designed by existing method. Along with the popularization of hybrid power filters, optimal design of parameters in the filters is an open problem.

Acknowledgements

This work is supported by Postdoctoral Science Foundation of China (2005037225).

References

1. Van Breemen, T.J., de Vries, A.: Design and Implementation of A Room Thermostat Using An Agent-based Approach. Control Engineering Practice 9, 233–248 (2001)
2. Zhao, S.G., Jiao, L.C., Wang, J.N., et al.: Adaptive Evolutionary Design of Passive Power Filters in Hybrid Power Filter Systems. Automation of Electric Power Systems 28, 54–57 (2004)
3. Zhu, X.R., Shi, X.C., Pend, Y.L., et al.: Simulated Annealing Based Multi-object Optimal Planning of Passive Power Filters. In: 2005 IEEE/PES Transmission and Distribution Conference & Exhibition, pp. 1–5 (2005)
4. Zhao, S.G., Du, Q., Liu, Z.P., et al.: UDT-based Multi-objective Evolutionary Design of Passive Power Filters of A Hybrid Power Filter System. In: Kang, L., Liu, Y., Zeng, S. (eds.) ICES 2007. LNCS, vol. 4684, pp. 309–318. Springer, Heidelberg (2007)
5. Lu, X.L., Zhou, L.W., Zhang, S.H., et al.: Multi-objective Optimal Design of Passive Power Filter. High Voltage Engineering 33, 177–182 (2007)
6. Lu, X.L., Zhang, S.H., Zhou, L.W.: Multi-objective Optimal Design and Experimental Simulation for Passive Power Filters. East China Electric Power 35, 303–306 (2007)
7. Zheng, Q., Lu, J.G.: Passive Filter Design Based on Non-dominated Sorting Genetic Algorithm II. Computer Measurement & Control 15, 135–137 (2007)
8. Carlos, A.C., Coello, R., Landa, B.: Evolutionary Multiobjective Optimization using A Cultural Algorithm. In: Proceedings of the Swarm Intelligence Symposium, pp. 6–13 (2003)
9. Guo, Y.N., Zhou, J.: Optimal Design of Passive Power Filters Based on Knowledge-based Chaotic Evolutionary Algorithm. In: Proceedings of the 4th International Conference on Natural Computation (accepted, 2008)

Robust Image Watermarking Scheme with General Regression Neural Network and FCM Algorithm

Li Jing[1,2], Fenlin Liu[1], and Bin Liu[1]

[1] ZhengZhou Information Science and Technology Institute , Zhengzhou 450002, China
[2] College of Information, Henan University of Finance and Economy
Jingli776@126.com, liufenlin@vip.sina.com, liubin@126.com

Abstract. A robust image watermarking method based on general regression neural network (GRNN) and fuzzy c-mean clustering algorithm (FCM) is proposed. In order to keep the balance between robustness and imperceptible, it uses FCM to adaptively identify watermarking embedding locations and strength based on four characteristic parameters of human visual system. For good learning ability and fast train speed of GRNN, it trains a GRNN with the feature vector based on the local correlation of digital image. Then embed and extracted watermark signal with the help of the trained GRNN. In watermark extracting it does not need original image. Experimental results show that the proposed method has better performance than the similar method in countering common image process, such as Jpeg compression, noise adding, filtering and so on.

Keywords: Image watermarking, general regression neural network, fuzzy c-mean clustering algorithm, human visual system, image processing.

1 Introduction

Digital watermarking is a technique that specified data (watermark) are embedded in a digital content. The embedded watermark can be detected or extracted later for ownership assertion or authentication. It has become a very active area of the multimedia security for being a way to counter problems like unauthorized handling, copying and reuse of information. However, it is required that the watermark is imperceptible and robust against attacks. In order to enhance the imperceptibility and robustness of watermark, recently the machine learning methods are being used in different watermarking stages such as embedding, detecting or extracting.

Yu et al. [1] proposed an adaptive way to decide the threshold by applying neural networks. Davis *et al.* [2] proposed a method based on neural networks to find maximum-strength watermarks for DWT coefficients. Knowles et al. [3] proposed a support vector machine aided watermarking method, in which the support vector machine is applied for the tamper detection in the watermark detection procedure. Fu *et al.* [4] have introduced support vector machine based watermark detection in the spatial domain, which shows good robustness. Shen *et al.* [6] employ support vector regression at the embedding stage for adaptively embedding the watermark in the blue channel in view of the human visual system.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 243 –250, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The above neural-network-based methods do enhance the robustness of watermarking techniques, but the natural limitations of the back-propagation learning algorithm in generalization capabilities and local minima counteract significant improvement. Support vector machine (SVM) has a better learning and generalization capabilities compared with the back-propagation learning algorithm. But the selection of SVM parameters is a difficult problem. Some people determine these parameters using cross-validation experiments. That is not fit watermarking because it takes up more time.

Inspired by the watermarking method in [6], we propose a watermarking scheme based on General Regression Neural Network (GRNN). GRNN is a powerful regression tool with a dynamic network structure. The network training speed is extremely fast and do not need to define more parameters by users. To keep the balance between imperceptibility and robustness, we use fuzzy c-mean (FCM) clustering algorithm to identify the watermark embedding location and embedding strength based on the human visual system (HVS).

This paper is organized in the following manner. The watermark inserting and detecting method is proposed in section 2; and Experimental results are shown in section 3; Section 4 is conclusions and analysis.

2 Watermarking Scheme

In an image, there is great correlation between any pixel value and the sum or variance of its adjacent pixels [5]. GRNN is used to learn the relationship. When the image undergoes some attacks, such as image processing, distortion, this relationship is unchanged or changes a little. So we can apply this feature for robust watermark embedding and extraction.

For human visual system, different location has different hiding ability in a natural image. Human visual system is more sensitive to the noise in smooth area than the texture area, and more sensitive to low luminance than high luminance. In general, high luminance and complex texture areas are fit to embed watermark signals. We use FCM clustering to locate embedding positions and decide embedding strength.

2.1 Locate the Embedding Positions with FCM

In the original gray image $I = \{g(i, j), 1 \le i \le M,\ 1 \le j \le N\}$, the perceptive characters of human visual system include 4 features: luminance sensitivity, texture sensitivity, contrast sensitivity and entropy sensitivity. These four measurements, as pixel point's feature vector, are regard as the sample data for FCM algorithm [7] for clustering, and divide pixel points into two clusters V_1 and V_2. The cluster V_1, whose members have larger feature values, is fit embedding watermark. The cluster V_2, whose members have smaller feature values, is not fit embedding watermark.

The basic idea of FCM algorithm is to establish a subject function for clustering, so that the distance inside the type is the shortest, that outside is the longest. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2 \tag{1}$$

where m is any real number greater than 1, x_i is the ith of d-dimensional sample data, c_j is the d-dimension center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center, u_{ij} is the degree of membership of x_i in the cluster j, called Subject degree, N is the number of sample data, C is the number of cluster. We set C=2 in this paper.

Given a pixel point (i, j) with the pixel value $g(i, j)$, its visual features are calculated using its eight neighboring pixel values $g(i-1, j-1)$, $g(i-1, j)$, $g(i-1, j+1)$), $g(i, j-1)$, $g(i, j+1)$, $g(i+1, j-1)$, $g(i+1, j)$, $g(i+1, j+1)$). The FCM algorithm is measured by the following four characteristics:

(1) Brightness sensitivity $B_{ij} = \sum_{m,n=-1}^{1} g(i+m, j+n)/9$

(2) Texture sensitivity $T_{ij} = \sum_{m,n=-1}^{1} | g(i+m, j+n) - B_{ij} |$

(3) Contrast sensitivity $C_{ij} = \max_{m,n=-1,0,1}(g(i+m,j+n)) - \min_{m,n=-1,0,1}(g(i+m,j+n))$

(4) Entropy sensitivity $E_{ij} = -\sum_{m,n=-1}^{1} p(i+m, j+n)\log p(i+m, j+n)$, where

$$p(i+m, j+n) = g(i+m, j+n) / \sum_{m,n=-1}^{1} g(i+m, j+n)$$

2.2 Watermark Image Pretreatment

In our scheme we choose a meaningful binary image to be the watermark. In order to resisting cropping and ensure the watermarking system security, the watermark image must be scrambled and scanned on to one-dimensional array.

For the watermark image $X=\{x(i, j), 1 \le i \le P,\ 1 \le j \le Q\}$, its pixel position reset according key K_1, then reshaped into one-dimensional sequence $W=\{w(i)\ \ i=1<i<P\times Q,,\ w(i) \in \{0,1\}\}$.

2.3 Train a GRNN

GRNN was presented by D. F. Specht in 1991 [7]. Based on nonlinear kernel regression in statistics, it can approximate the map inherent in any sample data, and the estimate can converge to the optimal regression surface even with sparse sample.

The GRNN includes 4 different layers: input layer, pattern layer, summation layer and output layer. The network input variables are $X = [x_1, x_2, \cdots, x_m]^T$ and output variables are $Y = [y_1, y_2, \cdots, y_l]^T$.

Compared with traditional back-propagation networks, GRNN includes the following advantages: (1) Fast training times; (2) Don't need determine the architecture of

the network. (3) Don't need to define some parameters such as learning rate, momentum and so on. But GRNN has trouble with irrelevant inputs, which is not a problem in our method because we extract the local correlative parameters to train a GRNN.

In this paper, GRNN is used to learn the relation between pixel value $g(i, j)$ and two feature values computed from its eight neighboring pixel values in a 3×3 window. The two feature values are pixel value sum S and pixel value variance D of eight neighboring pixel values, which is computed with the following formula:

$$S_r = \sum_{m,n=-1}^{1} g_r(i+m, j+n) - g_r(i, j) \tag{2}$$

$$D_r = \sum_{m,n=-1}^{1} [g_r(i+m, j+n) - \overline{g}_r]^2 - [g_r(i, j) - \overline{g}_r]^2 \qquad \overline{g}_r = S_r / 8 \tag{3}$$

So the input dataset is $X = \{S_r, D_r\}$, and the objective dataset is $Y = \{T_r \mid T_r = [g_r(i, j)]\}$. r is the number of samples.

Select L pixels from cluster *V1* randomly with the key *K2*, the L pixel values and their feature values $\{S_r, D_r, T_r\}$ $(r = 1,2,\cdots, L)$ as sample datasets to train a GRNN *net1*.

2.4 Watermark Embedding

Locate P×Q pixels positions from cluster *V1* randomly with key *K3* as watermark embedding positions. Note that the 3×3 windows centered these P×Q pixels are not overlapped.

Compute the feature values S and D of P*Q pixel as following vector:

$$P_r = \{S_r, D_r\} \qquad (r = 1,2,\cdots, P \times Q) \tag{4}$$

P_r is regarded as the input dataset of the trained GRNN *net1*, and get output vector:

$$\beta_r (r = 1,2,\cdots, P \times Q) \tag{5}$$

Referring to the output value, modify the pixel value $g_r(i, j)$ $(r = 1,2,\cdots, P \times Q)$ with the following formula (6), and get the watermarked image.

$$g'_r(i, j) = \begin{cases} \min(g_r(i, j), \beta_r - D \times u_r \times g_r(i, j)) \\ \qquad if \quad w(r) = 0 \\ \max(g_r(i, j), \beta_r + D \times u_r \times g_r(i, j)) \\ \qquad if \quad w(r) = 1 \end{cases} \tag{6}$$

where $g'_r(i, j)$ are the modified pixel values, D is the adjustable parameter, u_r is the Subject degree to the FCM cluster *V1* corresponding the *rth* pixel.

This embedding method modifies the pixel value in cluster *V1* that is not sensitive to human visual system, and changes embedding strength voluntarily based on the

Subject degree u_r, which is decided by the perceptive characters of human visual system from FCM. So this method is an adaptive watermarking algorithm that has stronger robustness.

2.5 Watermark Extracting

The watermark extraction is the reverse procedure of the watermark embedding. The detailed process is described as follows steps.

Step 1. With the key *K2* used in the watermark embedding process, select *L* pixels from the FCM cluster V_1' of test image I'. The *L* pixels and their feature values $\{S_r', D_r', T_r'\}$ $(r = 1,2,\cdots,L)$ as sample datasets to train a GRNN *net2*.

Step 2. With the key *K3* used in the watermark embedding process, select P×Q pixels from cluster V_1'. Compute the feature values of P×Q pixels to form a input vector $P_r' = \{S_r', D_r'\}(r = 1,2,\cdots,P\times Q)$, input the *net2* and get the output vector (7).

$$\beta_r'(r = 1,2,\cdots,P\times Q) \tag{7}$$

Step 3. Compare β_r' with the pixel value $g_r'(i,j)$, with the following formula extract the binary sequence $w'(r)$.

$$w'(r) = \begin{cases} 1 & if \quad g_r(i,j) > \beta_r' \\ 0 & otherwith \end{cases} \qquad r = 1,2,\cdots,P\times Q \tag{8}$$

Step 4. The extracted binary sequence $w'(r)$ is reshaped to P×Q two-dimensional matrix and then inversely scrambled with key *K1* used in watermark image pretreatment, get extracted binary image $X' = \{x'(i,j), 1 \le i \le P,\ 1 \le j \le Q\}$.

3 Experimental Results

In our experiments, the original images are 512×512 gray image Lena and Baboon, which have different texture, as show in fig. 1 (a) and (c). The input watermark is 58×28 binary image, as show in fig. 1 (e), and each pixel bit is as a watermark bit.

For the imperceptibility, a quantitative index, Peek Signal-to-Noise Ratio (PSNR), is employed to evaluate the difference between an original image *I* and a watermarked image I'. For the robust capability, the Bit Correct Rate (BCR) measures the bit difference between the original watermark $w(i)$ and the extracted watermark $w'(i)$. The PSNR is defined by formula (9)

$$PSNR = 10\times\log_{10}\frac{255^2}{MSE} \tag{9}$$

Where

$$MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y) - I'(x,y))^2}{M \times N} \tag{10}$$

A larger PSNR indicates that the watermarked image more closely resembles the original image, meaning that the watermarking method makes the watermark more imperceptible.

Figure 1 (b) and (d) are the watermarked Lena image and Baboon image, and their PSNR are 42.27 and 40.92 respectively. Figure 1 (e) is the original watermark image, Figure 1 (f) and (g) are the extracted watermark images from watermarked Lena image and watermarked Baboon image respectively. From this figure we see watermark can be extracted with no error bit from the un-attacked watermarked image.

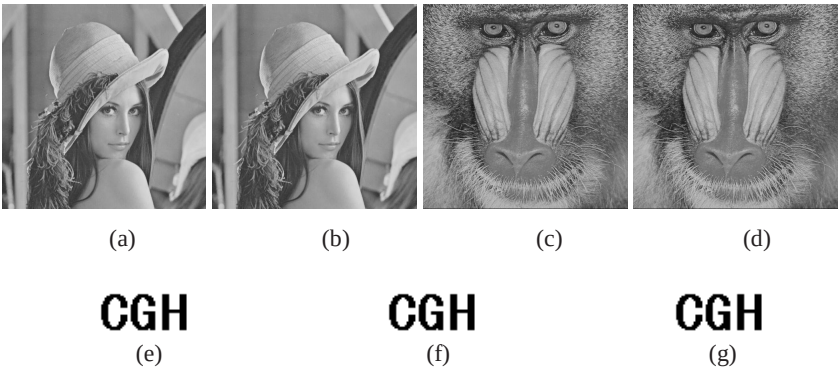


Fig. 1. (a) Original Lena image, (b) Watermarked Lena image (PSNR=42.27), (c) Original Baboon image, (d) Watermarked Baboon image (PSNR=40.92), (e) Original watermark image, (f) watermark extracted from (b) (BCR=1), (g) watermark extracted from (d) (BCR=1)

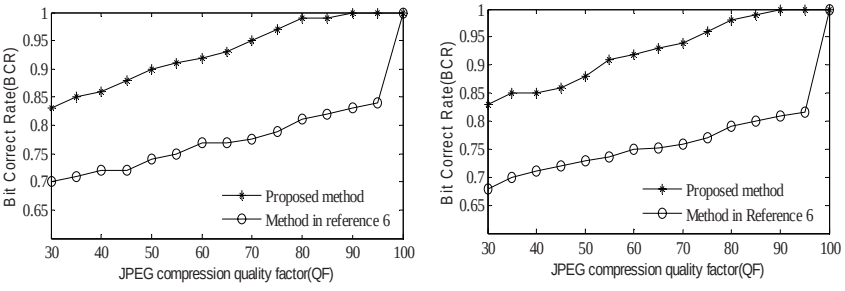
3.1 Robust against JPEG Compression

Although the proposed method embeds watermark in the space-domain, it holds stronger robustness to JPEG compression as frequency-domain methods. When the JPEG compression quality factor (QF) is larger than 90%, the watermark can be extracted with no error bit. When the QF is 30%, the extracted watermark image can be identified.

Figure 2 is the comparisons with the reference [6] on the attack of JPEG compressions. Figure 2 (a) and (b) shows the BCR of the watermark extracted from watermarked Lena image and Baboon image attacked by JPEG compressions with different QF respectively. From this figure we can see our method shows better performance to be against JPEG compressions compared with the method in reference [6].

3.2 Robust against Other Attacks

Experiments are also done to test robustness of the proposed method with respect to common attacks, including noise addition, image enhancement, filtering, cropping



(a) Test results in watermarked Lena (b) Test results in watermarked Baboon

Fig. 2. Robustness to JPEG compression

Table 1. Experimental results under different attacks compared with method in reference [6]

attacks	parameters	BCR of the extracted watermark			
		Watermarked Lena		Watermarked Baboon	
		Our method	Method in ref.6	Our method	Method in ref.6
Noise addition	Gaussian (0.05)	0.9347	0.9148	0.9287	60.919
	Salt and pepper(0.04)	0.9987	0.9573	0.9996	0.9486
	Speckle (0.04)	0.9732	0.9325	0.9831	0.9271
Image enhancement	Luminance (+50%)	0.9978	0.9728	0.9979	0.9637
	Contrast (+50%)	0.9997	0.9649	0.9982	0.9625
filtering	Low-pass filtering	1	0.9827	1	0.9853
	Median filtering	0.7363	0.6731	0.7126	0.5927
cropping	Cropping 1/4 image	0.9034	0.8934	0.9169	0.9102
blurring		0.9763	0.9324	0.9543	0.9428

and blurring. Table 1 lists BCR between original and extracted watermark from watermarked Lena and Baboon that have been attacked by common image processes. In order to show the comparing result, table 1 also lists the BCR gained by reference [6] method. From table 1 we can see our method outperforms the method in reference [6] in resisting image processing.

4 Conclusion and Analysis

This paper proposes a robust watermarking scheme based on general regression neural network and fuzzy c-mean clustering algorithm. In our work we find out:

(1) Combined with local correlation of digital image general regression neural network shows stronger function approximation and data predicting. Compared with SVR and back-propagation networks, GRNN does not need to set more parameters.

(2) GRNN has a large structure, which is a shortcoming for applying. But this is not important for the proposed method because it does not save trained neural net. It does not need original image in watermark detecting. That GRNN training speed is extremely fast is crucial for the proposed method.

(3) With the help of FCM, this method can voluntarily identify the watermark embedding locations and embedding strength. For this reason the proposed method has better performance than reference [6] in attacks, especially suffering from JPEG compression. At the same time we also find that the proposed method can not fight median filtering. The reason is that the neighbor pixel relations are destroyed when the image is processed by median filtering.

Acknowledgment. This work is supported partially by Science Technology Project of Henan, China (No.072102210001, No. 0624260019), National Natural Science Foundation of China (No.60774041)), and Natural Science Foundation of Education Department of Henan, China (No.2004922081).

References

1. Yu, P.T., Tsai, H.H., Lin, J.S.: Digital Watermarking Based on Neural Networks for Color Images. Signal Processing 81, 663–671 (2001)
2. Davis, K.J., Najarian, K.: Maximizing Strength of Digital Watermarks Using Neural Networks. In: International Joint Conference on Neural Networks, pp. 2893–2898. IEEE Press, New York (2001)
3. Knowles, H.D., Winne, D.A., Canagarajah, C.N., Bull, D.R.: Image Tamper Detection and Classification Using Support Vector Machines. In: IEEE Proceedings-Vision, Image and Signal Processing, pp. 322–328. IEEE Press, New York (2004)
4. Lu, W., Lu, H.T., Shen, R.M.: Color Image Watermarking Based on Neural Networks. In: Yin, F.-L., Wang, J., Guo, C. (eds.) ISNN 2004. LNCS, vol. 3174, pp. 651–656. Springer, Heidelberg (2004)
5. Sun, H.Z.: Contemporary digital image processing. Electron and Industry Publishing Company, Beijing (2006)
6. Shen, R., Fu, Y., Lu, H.: A Novel Image Watermarking Scheme Based on Support Vector Regression. Journal of System and Software 78(1), 1–8 (2005)
7. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York (1981)
8. Specht, D.F.: A General Regression Neural Network. IEEE Transactions on Neural Networks 2(6), 68–576 (1991)

Research on Sampling Methods in Particle Filtering Based upon Microstructure of State Variable

Zhiquan Feng[1], Bo Yang[1], Yuehui Chen[1], Yanwei Zheng[1], Yi Li[2], and Zhonghua Wang[2]

[1] School of Information Science and Engineering, University of Jinan, Jinan, 250022, P.R. China
[2] School of Control Science and Engineering, University of Jinan, Jinan, 250022, P.R. China
ise_fengzq@ujn.edu.cn

Abstract. With the purpose of decreasing the number of particles needed to be sampled stochastically in particle filtering, a new sampling method used in particle filtering is put forward in this paper. First of all, under specific human-computer conditions, both cognitive psychology features of operators and motion features of the operators' hands are studied, upon which a novel concept, microstructure of state variable (MSV), is proposed. Then, a general method to describe the MSV is further discussed. At last, a Sampling Algorithm based on the MSV is put forward, which is also called SAMSV. The MSV provides an unique and efficient mathematic model that makes SAMSV avoid sampling huge particles with poor quality. In order to demonstrate validity and performance of SAMSV, a great deal of experiments is completed. By using fewer particles, compared with conventional particle filtering, the SAMSV may achieve better tracking precision. In forms of both the theoretical analysis and a great deal of experimental results from real video data, just a few particles are needed to describe post probability distribution of state variables without reducing the tracking precision by our dimensionality reduction method.

Keywords: Human hand tracking, dimensionality reduction method, Human computer interaction.

1 Introduction

3D human hand gesture tracking is one of the kernel contents studied in human-computer interface. One of the main methods to implement hand gesture tracking is particle filtering(PF)[1]. PF is an effective approach to deal with non-Gaussian and non-linear problems, which was developped from the middle of 1990s, and it has been regarded as one of the most promising methods to estimate states so far. PF needs a great deal of samples to describe posteriori distribution of a hand gesture state, which makes real-time tracking impossible. For example, for a state with 10 dimension variables, and takes just 20 dispersive values in each dimension, it is clear that 320,000,000 examples are required to only cover the typical set of posteriori distribution[2]. So, how to decrease the dimensions of the state space is an intractable problem, and has been widely studied for many years. Some theories have been proposed by now.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 251–257, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The first method is to decompose state space into sub-spaces[3-7], and then PF is imposed respectively on each sub-spaces. As a substitute method, searches are processed in each sub-space, the results are used as restrictions upon which search space of the left variables is reduced. It is difficult to subdivide the state space, and even if this decomposition is implemented, it is not appropriate to deal with them independently[8].

The second method is to obtain the relationship between hand gestures and high-dimensionality feature vector sets by means of machine learning[9-10]. One of the main characteristics of the method is that 3D hand gesture estimation is regarded as a procedure of learning and reasoning.

The third method uses the technique of analysis-by-synthesis and coarse-to-fine to deal with high-dimensionality problem[11-16]. In this approach, fitting image data to the hand gesture models is complex and would be quite slow.

PF is combined with constrained variable metric gradient descent methods in M. de La Gorce and N. Paragios's studies[17], and convergence rate is greatly improved.

Aiming at describing post probability distribution of state variables by using a small amount of particles, a new sampling approach in PF is put forward in this paper.

2 Microstructure of State Variable (MSV)

2.1 Core Ideas

An effective and novel study idea is to discuss sampling methods under specific human-computer interaction. Our research is under the background of HCI in a virtual prototyping system, both cognitive psychology features of operators and motion features of operators' hands are used as the breakthrough points. By means of decreasing particle numbers needed to be sampled stochastically, the approaches to sample particles are studied systemically and globally in this paper.

2.2 Experimental Scheme

According to contemporary cognitive psychology[18], in the process of human cognitive activity, when human cognitive activities occur, events or information is dealt with one by one, with sequence and continuity. On the other hand, according to human body athletics physiology[19], human body muscle takes on two physiological features: excitability and contractibility.

Based upon the above basic theories, a complete set of experimental scheme is devised in order to study both cognitive psychology features of operators and motion features of the operators' hands. All experimenters are divided into five groups in advance, with ten experimenters each group, and those who belong to the same group accomplish the same typical operation. At the same time, the experimenters are equipped with data gloves on their right hands, data from sensors is transmitted into a computer. Fig.1 shows some of the typical operations, in which some snippets are displayed in each row.

Fig. 1. Some of the typical operations devised for exploring both cognitive psychology features of operators and motion features of the operators' hands

2.3 Experimental Results

5000 consecutive frames are analyzed in each typical operation, and each joint angle changing in every frame are studied, focusing on exploring changes of speed and acceleration and angles. Our statistical number shows that the number of states that the ratio of the frame number with speed kept unchanging to the total number, 5000, is 59.26%, on average, and that the number of states that the ratio of the frame number with acceleration kept unchanging to the total number, 5000, is 52.81%, on average.

2.4 Concept of MSV

MSV could be described by using the four channels, which are shown in Fig.2. In Fig2, S_t expresses the value of a state variable at time t. The Even Speed Channel (ESC) is used to depict features involved into even speed motion of a state variable, and the Even Acceleration Channel (EAC) is used to depict features which are involved into even accelerated motion of a state variable. It is very clear that, at time t, a state variable must choose one and only one channel to enter. No matter the ESC or the EAC, the way to enter them are based on randomized algorithms. Real value of

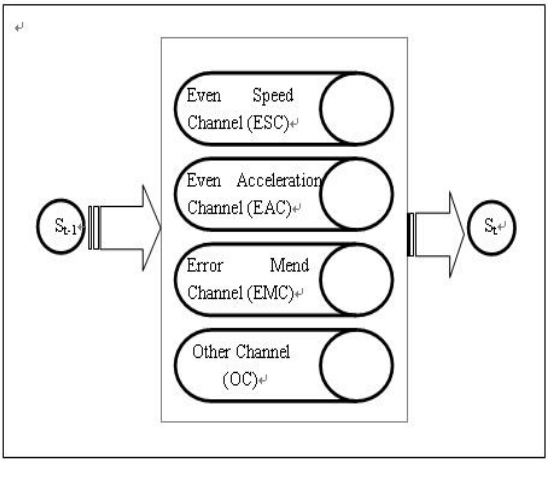


Fig. 2. The structure of MSV

the variable is replaced with its predicted value in our algorithm, which makes it possible that the estimation value of the variable gets away from its real value. So, a new channel, called the Error Mend Channel (EMC), is introduced into MSV. The states which fail to enter ESC, EAC or EMC have to enter into OC.

The MSV has an important characteristics. Let a state be composed of n variables, and $X=M_1 \cup M_2$, among which, M_1 is a set in which variable values are computed by means of deterministical methods with huge probability, and M_2 is a set in which variable values are computed by means of randomized methods with huge probability, then, the process to describe typical set of X be equal to that to describe typical set of M_2.

3 Sampling Algorithm Based on MSV(SAMSV)

All channels are divided into three priorities, among which, the EMC takes on the highest priorities of all channels, the EAC and the EAC possess inferior priorities, and the OC is provided with the lowest priority. In other words, the EMC is judged and processed firstly.

The new sampling method based on the MSV is described as follows:

Step 1: All of the channel probabilities are computed in terms of formula (1)—(4).

Step 2: All of the channel probabilities are united, obtaining united probabilities $P^{(1)}$, $P^{(2)}$, $P^{(3)}$ and $P^{(4)}$, or

$$\sum_{k=1}^{4} P^{(k)} = 1$$

Step 3: If

$$P_3 > T_e$$

enter into EMC.

Step 4: According to the united probabilities, channels are selected based upon probability sampling, and then different methods are chosen in terms of the channels selected.

Case ESC: The current state variable value is computed in terms of even speed motion models.

Case EAC: The current state variable value is computed in terms of even accelerated motion models.

Case OC: The value of the current state variables is obtained by means of sampling stochastically based on the prior probability distributions.

Step 5: Loop between Step 3 and Step 5, until all variables in the state vector are gained.

Step 6: Return the particle.

This algorithm is also called SAMSV in this paper. It should be pointed out that, both the OC and the EAC have the same processing method, but they have different priorities. The stochastic sampling is carried on along the variable gradient directions.

4 Algorithm Analysis

The time cost in computing channel probabilities depends mainly on finding local error, Δ_{t-1}, under the worst situation, the errors between 20 projected points of joints and corresponding frame image contour are needed to be computed, and so $20d_0$ loops are required, which d_0 is the maximum error of the all errors. Let N be the particle number, it is clear that the total time complexity is $O(N*d_0)$.

5 Experimental Results

An operator moves her hand without any marks on it under a single camera, and frames with 60ms interval between consecutive frames are used as input data of our experiment, in which backgrounds are segmented by manpowers. The experiment is divided into two parts. The one is to implement PF based on the MSM (PF-MSM) put forward in this paper, and the another one is to realize the classical PF[17] for comparison with the first part.

First of all, the PF-MSM is implemented. The experimental results are presented in Fig.3. Fig.3(a) are the tracked 3D hand gesture models to the same frame, with different particles, 100, 2000, and 3000, respectively, from top to bottom and from left to right. Fig.3(b) are the projections of the corresponding 3D hand gesture models onto the corresponding frame images.

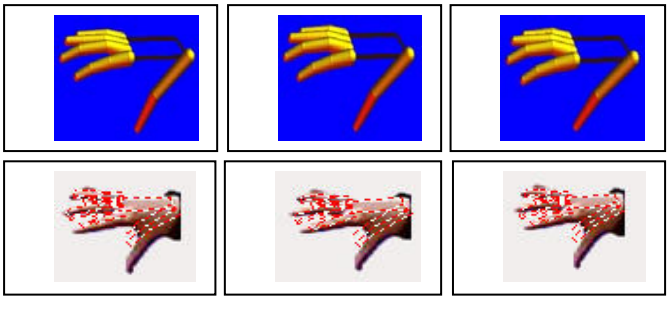


Fig. 3. The tracked results using PF based on SAMSV, with different particles

It is demonstrated from Fig.3 that SAMSV is feasible and effective.

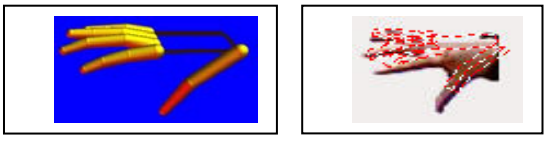


Fig. 4. Tracking results using classical PF [17], with 1000 particles

In order to demonstrate the tracking precision and performance of SAMSV, the referenced classical PF[17] is used as the referenced method. On the case of 1000 particles, experimental results are presented in Fig.4. Fig.4(a) is the tracked 3D hand models, and Fig.4(b) is the projection of the 3D hand gesture model onto the corresponding frame image.

On personal computer with Pentium 4 and 1.80G HZ CPU and 256MB memory, the average running time of PF-SAMSV and the classical PF are 276.43 millisecond per frame (ms/per frame) and 345.79 ms/per frame, and the running time of PF-SAMSV is decreased by 20.06% compared with that of the classical PF on the average within 100 particles.

It can be shown from Fig.4 that the effect under the condition of 1000 particles is equivalent to that under the condition of 20 particles in Fig.3.

6 Future Work and Conclusions

The main contributions of this paper are as followings: (1) it is encouragingly perspective to use both cognitive psychology features of operators and motion features of the operators' hands for the high-dimensionality sampling problem. (2) the MSV based on the cognitive psychology features of operators and motion features of the operators' hands could describe stochastically change trends of state variables with comprehensiveness, delicateness and uniform, which provides an effective, efficient and uniform mathematic model and data structure for design of sampling algorithms.

Acknowledgements

This paper is supported by National Natural Science Foundation of China(60773109), Natural Science Foundation of Shandong Province(Y2007G39), Key Project of Natural Science Foundation of Shandong Province(2006G03), Science and Technology Plan of Shandong Province Education Department(J07YJ18), PhD Scientific Research Foundation(XBS0705).

References

1. Gordon, N., Salmond, D.J., Smith, A.F.M.: Novel Approach to Nonlinear and Non-Gaussian Bayesian State Estimation [J]. IEE Proceedings-F 140, 107–113 (1993)
2. Rui, C., Guoyi, L., Guoying, Z., Jun, Z., Hua, L.: 3D Human Motion Tracking Based on Sequential Monte Carlo Method. Journal of Computer Aided Design & Computer Graphics 17(1), 85–92 (2005)
3. MacCormick, J., Isard, M.: Partitioned Sampling, Aarticulated Objects, and Interface-Quality Hand Tracking. In: Vernon, D. (ed.) ECCV 2000. LNCS, vol. 1843, pp. 3–19. Springer, Heidelberg (2000)
4. Deutscher, J., Davidson, A.: ReidI1.: Articulated Partitioning of High Dimensional Search Spaces Associated With Articulated Body Motion Capture[C]. In: Proceedings of International Conference on Computer Vision and Pattern Recognition, Hawaii, vol. 2, pp. 669–676 (2001)

5. MacCormick, J., Blake, A.: Partitioned Sampling, Articulated Objects and Interface-quality Hand Tracking[C]. In: Proc. 7th European Conf. Computer Vision, Dublin, Ireland, pp. 3–19 (2000)
6. Wu, Y., Huang, T.S.: Capturing Articulated Human Motion: A Divide-and-Conquer Approach[C]. In: Proc. IEEE Int. Conf. Computer Vision, pp. 606–611 (1999)
7. Zhou, H., Huang, T.S.: Tracking Articulated Hand Motion with Eigen Dynamics Analysis[C]. In: Intl. Conf. on Computer Vision, Nice, France, pp. 1102–1109 (2003)
8. Deutscher, J., Davison, A., Reid, I.: Automatic Partitioning of High Dimensional Search Spaces Associated With Articulated Body Motion Capture. In: Proc. Conf. Computer Vision and Pattern Recognition, pp. 187–193 (2001)
9. Wachs, J.P., Stem, H., Edan, Y.: Real-Time Hand Gesture Telerobotic System Using the Fuzzy C-Means Clustering Algorithm[C]. In: World Automation Congress(WAC), Orlando, Florida, USA, pp. 403–410 (2002)
10. Rosales, R.: The Specialized Mappings Architecture with Applications to Vision-Based Estimation of Articulated Body Pose[D]. Ph.D. Thesis.Boston University Graduate School of Arts and And Sciences (2002)
11. Rehg, J.M., Kanade, T.: Visual Tracking of High DOF Articulated Structures: An Application to Human Tracking[C]. In: Eklundh, J.-O. (ed.) Proc. 3rd European Conference on Computer Vision, Sweden, vol. 801, pp. 35–46 (1994)
12. Rehg, J.M., Kanade, T.: Digiteyes: Vision-Based Human Hand Tracking[R]. Tech. Rep. CMU-CS-93-220, School of Computer Science, Carnegie Mellon University (1993)
13. Lee, S.U., Cohen, I.: 3D Hand Reconstruction From Monocular View. International Conference on Pattern Recognition[C]. In: Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004), Cambridge, UK, vol. 3, pp. 310–313 (2004)
14. Lin, E., Cassidy, A., Hook, D., Baliga, A., Chen, T.: Hand Tracking Using Spatial Gesture Modeling and Visual Feedback for a Virtual DJ System[C]. In: Fourth IEEE International Conference on Multimodal Interfaces (ICMI 2002), NY, USA, vol. 2, pp. 197–202 (2002)
15. Pan, C., Ma, S.: Region-Based 3D Motion Tracking of Hand[J]. Chinese Journal of Image and Graphics 8(10), 1205–1210 (2003)
16. Cui, J.: Studies on Three-Dimensional Model Based Posture Estimation and Tracking of Articulated Objects[D]. Doctor degree dissertation, Tsinghua University, Beijing, China (2004)
17. Gorce, M., Paragios, N.: Monocular Hand Pose Estimation Using Variable Metric Gradient-Descent. In: British Machine Vision Conference, British Machine Vision Conference (BMVC), Edinburgh, vol. 3:1, pp. 269–1278 (2006)
18. Liang, J.: Contemporary Cognitive Psychology[M]. Shanghai Education Publication House, Shanghai (2006)
19. Cao, Z.: Athletics Physiology[M]. Higher Education Press, Beijing (1999)

Knowledge-Supported Segmentation and Semantic Contents Extraction from MPEG Videos for Highlight-Based Annotation, Indexing and Retrieval

Jinchang Ren, Juan Chen, Jianmin Jiang, and Stan S. Ipson

School of Informatics, University of Bradford, BD7 1DP, UK
{j.ren,j.chen12,j.jiang1,s.s.ipson}@bradford.ac.uk
http://dmsri.inf.brad.ac.uk/

Abstract. Automatic recognition of highlights from videos is a fundamental and challenging problem for content-based indexing and retrieval applications. In this paper, we propose techniques to solve this problem by using knowledge supported extraction of semantic contents, and compressed-domain processing is employed for efficiency. Firstly, video shots are detected by using knowledge-supported rules. Then, human objects are detected via statistical skin detection. Meanwhile, camera motion like zoom in is identified. Finally, highlights of zooming in human objects are extracted and used for annotation, indexing and retrieval of the whole videos. Results from large data of test videos have demonstrated the accuracy and robustness of the proposed techniques.

Keywords: Video segmentation, highlights extraction, skin detection, camera motion, MPEG, shot detection.

1 Introduction

With the development of network and multimedia techniques, content-based analysis, indexing and retrieval of videos has attracted numerous attentions for its broad applications and huge commercial potential, such as digital library, video on demand, telemedicine, etc [1, 2, 6]. Among these applications, a fundamental problem is how to extract meaningful highlights from videos, which is highly desired as it can help to automatic annotate the video contents and abstract the whole video for fast browsing and searching. Unfortunately, this problem is still far away from completely resolved towards real applications.

In general, videos are usually segmented into short clips namely shots and scenes, etc. A video shot contains frames from a certain camera motion, such as zooming, panning or tilting, and different shots are composed together by some special editing effect like fade in/fade out, dissolve, etc [1-4]. As for video scene, usually it contains several shots to represent a relatively complete semantic content. Moreover, people also introduce key frames to characterize each shot or scene. Consequently, there are at least 4 levels in hierarchical representation of video, including key frames, shots, scenes and the whole video. Although this hierarchical structure provides a practical way for video representation, it lacks semantic contents to match the requirements

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 258–265, 2008.
© Springer-Verlag Berlin Heidelberg 2008

from general users. Therefore, it is our aim to extract highlights from videos and annotate the video shots for further highlights-based indexing and retrieval.

In most of the videos, human activity always is the focus. Detection of skin pixels in videos plays important roles for many relevant image and vision applications, such as face detection, facial expression recognition, gesture recognition, human-computer interaction and even naked people detection [12-16]. In many video applications, highlights are emphasized by zooming-in of human objects, and these can be widely found in TV news, sports games, and films, etc [5-11]. As a result, we intend to detect this scenario in videos for intelligent annotation and retrieval of video.

The main contributions of this paper can be summarized below. Firstly, we proposed effective ways in feature extraction from DC images and motion vectors, which are then successfully employed for video shot detection. Secondly, these compressed domain features are applied in estimation of camera motion. Then, statistical modeling is used for skin detection. Finally, an effective solution is introduced to detect highlight of zooming-in human objects for automatic annotation, indexing and retrieval of video. Since all the relevant techniques are based on compressed MPEG videos, our system can achieve real-time performance for online applications.

The rest of the paper is organized as follows. Section 2 gives a comprehensive literature review in terms of video segmentation, skin detection and zooming-in detection. Based on which, Section 3 discusses the details on algorithm design of the proposed system including feature extraction, video segmentation, statistical skin detection, camera motion classification and finally detection of zooming-in human objects. Experimental results and discussions are illustrated in Section 4 with brief conclusions and acknowledgement given in Section 5 and Section 6, respectively.

2 Literature Review

In this section, relevant work in terms of video segmentation, skin detection and camera motion detection is reviewed. Among these techniques developed, we are focusing on knowledge-supported ones towards extraction of semantics and highlights for video indexing and retrieval. In addition, to lower the computational cost in decoding the whole sequence, compressed domain processing is employed for efficiency as most of the media data is now available in compressed formats.

Generally, domain knowledge is very essential in extraction of highlights and automatic annotation of the videos, as semantic contents within the videos are normally context-based. For example, highlights in sports videos like closed captions, slow-motion replays and special zoom [5-11] are quite different from those of political news. However, it is found that almost all these highlights have corresponding camera zoom in motion, i.e. close-up events. According to statistical results from [11], such highlight of player close-up occupies 39.1%, 25.1%, 39.4%, 48.1% and 57.6% in their predefined shot classes in tennis, soccer, basketball, volleyball and table tennis games, respectively. Therefore, detection of such events seems a straightforward solution in such a context. Relevant investigations in video segmentation, camera motion detection and skin pixel classification are discussed below, respectively.

As for shot boundary detection, it was originally introduced decades ago to detect abrupt cut in videos [1-4]. From then on, many techniques have been developed in either compressed domain or uncompressed domain. With features extracted from compressed and/or uncompressed domain, a continuity signal is constructed which represents similarity between neighboring frames. As for decision, rule-based approaches and statistical machine learning like SVM and neural network become more popular than traditional thresholding based approaches.

Camera motion is important in at least two aspects: i) it can help to reduce false alarms in shot detection; ii) it can be taken for event detection and automatic annotation of video contents. In most of cases, only pan, tilt and zoom factors are considered in such estimations [3-5]. Since motion vectors have been extracted in compressed videos, compressed-domain processing to estimate camera motion seems more efficient. In Zhang et al [8], sum of difference between motion vectors and change of signs of motion vectors (across the zoom center) are used in estimating pan, tilt and zooming factors. In Kobla el at [19], dominant motion in directional histogram of 8 bins is taken to estimate pan and tilt, and then focus of contraction and focus of expansion are used in determining zoom parameter. In [20], arbitrary camera motion is estimated from MPEG videos by removing outlier motion vectors. In [4], Tan et al introduced rapid estimate of camera motion from P-frames in MPEG videos, however, it suffers inaccuracy caused by large moving objects. Consequently, an improved version is introduced in Section 3.3 for more robust detection of camera motions.

Regarding human detection, classification of skin colors is widely employed. Although many color spaces have been applied in skin detection, including RGB [14], HSV [13], YUV/YCbCr [12], etc., they can be simply categorized into two classes by examining whether the luminance intensity component is utilized.

3 Algorithm Design

In our system, the input data stream is MPEG compressed videos. It is well-known that block-based motion estimation and data representation form the fundamental structure in MPEG standard, and a macroblock of 16×16 pixels is the basic element in analysis and coding as motion vectors are defined on these macroblocks. In addition, each macroblock can be further divided into four 8×8 sub-blocks, based on which DCT and IDCT are implemented. Hence there are one DC coefficient and 63 AC coefficients corresponding to every sub-block. For each macroblock, we have 4 luminance sub-blocks but one chromatic sub-block for either Cb or Cr component when 4:2:0 chromatic representation is used. Further details on compressed-domain feature extraction, video segmentation and highlights detection on the basis of camera motion estimation and skin detection are discussed in the following sections.

3.1 Feature Extraction in Compressed Videos

For the i^{th} input frame f_i, let us define N_h and N_v denote number of macroblocks in horizontal and vertical directions. Let (n_x, n_y) denote the position of one

macroblock in frame i, where $n_h \in [1, N_h]$ and $n_v \in [1, N_v]$; and $Y_{dc}(i), U_{dc}(i)$ and $V_{dc}(i)$ represent the corresponding DC images of Y, Cb and Cr component.

Since all the macroblocks are intra-coded in I-frames, the corresponding DC images can be directly extracted. As for P-frame and B-frame, weighted motion compensation is applied as the current macroblock may contain contributions from its four original neighboring blocks in the reference frame. It is easily proved that each DC value equals 8 times of the average intensity in the related subblock, hence the DC image provides a low-resolution version of the original image for analysis.

Denote N_y, N_{cb} and N_{cr} as numbers of elements in $Y_{dc}(i), U_{dc}(i)$ and $V_{dc}(i)$, respectively. Considering 4:2:0 chromatic format, we have $N_{cb} = N_{cr} = N_h N_v$ and $N_y = 4N_h N_v$. For frame i and j, its normalized luminance difference is defined as

$$D_y(i, j) = (8N_y \cdot 255)^{-1} \sum_{n=1}^{N_y} \left| Y_{dc}(i,n) - Y_{dc}(j,n) \right| \tag{1}$$

Similarly, normalized differences in two chromatic components can be defined as $D_u(i, j)$ and $D_v(i, j)$, respectively . Moreover, we define M_i as the motion magnitude measurement in frame i, on the basis of extracted motion vectors $V_x(i,n)$ and $V_y(i,n)$ from the n^{th} macroblock.

$$M_i = \frac{1}{\alpha_i} \sum_{n=1}^{N_h N_v} \beta_n \left(\left| V_x(i,n) \right| + \left| V_y(i,n) \right| \right), \quad \beta_n = \begin{cases} 1 & \text{if} \quad \left| V_x(i,n) \right| + \left| V_y(i,n) \right| \geq 2\overline{M}_i / 3 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$M_i = (N_h N_v)^{-1} \sum_{n=1}^{N_h N_v} \left[\left| V_x(i,n) \right| + \left| V_y(i,n) \right| \right], \quad \lambda_i = \alpha_i (N_h N_v)^{-1} \tag{3}$$

.

where α_i is the total number of valid motion vectors with their sum of absolute velocities in x/y directions larger than an adaptive threshold of $2\overline{M}_i / 3$. Moreover, λ_i indicates a ratio of verified motion vectors in the whole frame.

3.2 Shot Detection for Video Segmentation

According to our observations from over 25,000 frames of sequences containing more than 150 shots, it is found that shot boundary usually contain apparent luminance and chromatic difference. However, there are also some false alarms introduced by either local motion or flicker/flashing light. Hence our method is designed below:

i) Firstly, each frame f_i is compared with its previous one, f_{i-1}, in terms of D_y, D_u and D_v against three thresholds y_{th}, u_{th}, v_{th}. If either of them exceed the corresponding threshold, a candidate shot change is detected;

ii) To verify each of the candidate shot changes, a sliding window of 3 frames is introduced, in which f_{i-1} and f_{i+1} are compared. Only if the corresponding difference of D_y, D_u or D_v exceeds y_{th}, u_{th} or v_{th}, the shot change is verified. Otherwise, it is a false alarm caused by either flicker or flashing light;

iii) To cope with motion effect, the thresholds of y_{th}, u_{th} and v_{th} are not fixed. Instead, their values are selected to adapt with the image contents in the videos as discussed in Eq. (7) below.

$$y_{th} = y_0 \lambda_i;\ u_{th} = u_0 \lambda_i;\ v_{th} = v_0 \lambda_i . \tag{4}$$

In (4), y_0, u_0 and v_0 are constants which are further used to decide y_{th}, u_{th} and v_{th} by considering λ_i, the ratio of verified motion vectors in f_i defined in (3).

3.3 Improved Estimation of Camera Motion

According to the 6-parameter projective camera model only considering rotation and zoom between frames in (5), it is found in [4] that p_1 indicates camera zoom factor ($p_1 > 1$ represents zoom in and $p_1 > 1$ represents zoom out), where (x_i, y_i) and (x_{i-1}, y_{i-1}) are the image coordinates of corresponding points in f_i and f_{i-1}.

$$x_i = \frac{p_1 x_{i-1} + p_2 y_{i-1} + p_3}{p_5 x_{i-1} + p_6 y_{i-1} + 1}; \qquad y_i = \frac{-p_2 x_{i-1} + p_1 y_{i-1} + p_4}{p_5 x_{i-1} + p_6 y_{i-1} + 1}. \tag{5}$$

As we pointed out in Section 2, the above solution suffers false alarms caused by object motion. In our improved approach, the method introduced in [21] is employed to remove outlier motion vectors. Hence not all the inter-coded macroblocks are used in estimating of camera motion. Instead, only macroblocks with their motion vectors satisfying smoothness conditions in [21] are considered to estimate p_1 .

3.4 Skin Detection

YCbCr color space is utilized in our approach as it is easily extracted from MPEG data. For each labeled block of skin colors, its associated color entry is denoted as $c_e = (y, c_b, c_r)$, then the corresponding histogram is attained in which probability or likelihood of a color entry representing skin is determined below,

$$p(c_e / skin) = sum(c_e / skin) / N_s . \tag{6}$$

where $sum(c_e / skin)$ denote number of occurrence in training data when the color entry c_e is defined as skin, and N_s indicates total number of occurrences.

In our system, non-skin colors are also trained with their likelihood as skin near or equaling to zero. With the extracted statistical model of skin colors, probability or likelihood of skin of each color entry c_e can be easily determined by checking

$p(c_e / skin)$ as a look-up table. A threshold L_{th} is then introduced for classification, which is decided by ROC analysis of the skin model on a tradeoff between correct detection rate R_c and false positive rate R_f.

$$skin(b, c_e) = \begin{cases} true & if \quad p(c_e / skin) > L_{th} \\ false & otherwise \end{cases}. \tag{7}$$

Through analysis of ROC curves of skin detection results from trained data by using YCbCr and CbCr color spaces, we have found that YCbCr space yields better performance than CbCr space, although some others believed that the latter should be more robust [16]. The threshold L_{th} is chosen when $R_c = 90\%$ and $R_f = 4.2\%$.

3.5 Overall Workflow for Video Highlights Extraction

The overall workflow of our system for video highlights extraction is given below: Firstly, detect shots from the input video, and in each shot, determine frames with camera zoom-in. Then, detect skin templates from these zoom-in frames and these are then taken as highlights of zooming-in human objects. Finally, these extracted highlights are further used as semantic concepts in video annotation and retrieval.

4 Results and Discussions

In our experiments, we have 6 test sequences as summarized in Table 1 which contains three categories, i.e. film, sports and news. In total there are 68225 frames segmented into 227 shots, and we also manually labeled 18 highlights of zooming in human objects from these videos. Examples of sample images from two test sequences and associated skin pixels detected are shown in Fig. 1.

Results on shot detection and highlights extraction from each test sequence are listed in Table 2 for comparisons. From Table 2 we can find several facts below:

i) Our shot detection approach yield quite promising results of good precision rate and recall rate;
ii) As for highlights extraction, our recall rate is fairly good but precision needs to be improved. The reason is the video contains only a few highlights hence even one false alarm may cause poor precision rate in quantitative evaluation;
iii) With extracted highlights, original video data is then annotated and indexed by using this semantic concept for more effective retrieval.

Table 1. Details information about our test sequences

Sequence	(a)	(b)	(c)	(d)	(e)	(f)	Sum
Category	Film	Film	Sports	Sports	News	News	N/A
Frames	13281	6900	9605	22162	7554	8723	68225
shots	57	29	31	76	21	23	227
Highlights	3	2	3	5	2	3	18

Fig. 1. Two sample images (left) and detected blocks of skin pixels

Table 2. Performance in terms of shot detection and highlights extraction

Sequence		(a)	(b)	(c)	(d)	(e)	(f)	Average
Shot Detection	Precision	83.3%	80.2%	84.9%	84.2%	81.1%	79.1%	81.7%
	Recall	91.5%	92.0%	93.8%	92.4%	90.2%	87.6%	91.2%
Highlights extraction	Precision	67.7%	67.7%	75.0%	80.0%	100.0 %	75.0%	77.6%
	Recall	67.7%	100.0%	100.0%	80.0%	100.0 %	100.0%	91.3%

5 Conclusions

We presented an approach for extraction of highlights in compressed videos. We found that features from DCT domain and YUV color spaces from MPEG videos yielded quite good performance in shot detection, camera motion determination and statistical skin detection. Our experiments from a variety of test sequences have successfully demonstrated the effectiveness of the proposed techniques. As the whole system is implemented in compressed domain, this will enable real-time processing for many potential online applications.

Acknowledgement. Finally, the authors wish to acknowledge the financial support under EU IST FP-6 Research Programme with the integrated project: LIVE (Contract No. IST-4-027312).

References

1. Cotsaces, C., Nikolaidis, N., Pitas, I.: Video Shot Detection and Condensed Representation: A Review. IEEE Signal Proc. Mag. 23(2), 28–37 (2006)
2. Rasheed, Z., Shah, M.: Detection and Representation of Scenes in Videos. IEEE T-Multimedia 7(6), 1097–1105 (2005)
3. Fang, H., Jiang, J., Feng, Y.: A Fuzzy Logic Approach for Detection of Video Shot Boundaries. Pattern Recognition 39(11), 2092–2100 (2006)
4. Tan, Y.P., Saur, D.D., Kulkarni, S.R., Ramadge, P.J.: Rapid Estimation of Camera Motion from Compressed Video with Application to Video Annotation. IEEE-TCSVT 10(1), 133–146 (2000)
5. Yuan, J., Wang, H., Xiao, L., Zheng, W., Li, J., Lin, F., Zhang, B.: A Formal Study of Shot Boundary Detection. IEEE-TCSVT 17(2), 168–186 (2007)

6. Sebe, N., Lew, M.S., Smeulders, A.W.M.: Editorial Introduction: Video retrieval and Summarization. J. CVIU. 92(2-3), 141–146 (2003)
7. Boccignone, G., Chianese, A., Moscato, V., Picariello, A.: Foveated Shot Detection for Video Segmentation. IEEE T-CSVT 15(3), 365–377 (2005)
8. Zhang, H.J., Low, S.Y., Smoliar, S.W.: Video Parsing and Browsing using Compressed Data. Multimedia Tools and Applications 1(1), 89–111 (1995)
9. Bescos, J., Cisneros, G., Martinez, J.M., Menendez, J.M., Cabrera, J.: A Unifoed Model for Techniques on Video-Shot Transition Detection. IEEE T-Multimedia 7(2), 293–307 (2005)
10. Ekin, A., Tekalp, M., Mehrotra, R.: Automatic Soccer Video Analysis and Summarization. IEEE T-IP 12(7), 796–807 (2003)
11. Duan, L.-Y., Xu, M., Tian, Q., Xu, C.-S., Jin, J.S.: A Unified Framework for Semantic Shot Classification in Sports Video. IEEE T-Multimedia 7(6), 1066–1083 (2005)
12. Phung, S.L., Bouzerdoum, A., Chai, D.: Skin Segmentation Using Color Pixel Classification: Analysis and Comparison. IEEE T-PAMI 27(1), 148–154 (2005)
13. Sigal, L., Sclaroff, S., Athitsos, V.: Skin Color-Based Video Segmentation under Time-Varying Illumination. IEEE T-PAMI 26(7), 862–877 (2004)
14. Jones, M.J., Rehg, J.M.: Statistical Color Models with Application to Skin Detection. Int. J. Computer Vision. 46(1), 81–96 (2002)
15. Kakumanu, P., Makrogiannis, S., Bourbakis, N.: A Survey of Skin-Color Modeling and Detection Methods. Pattern Recognition 40, 1106–1122 (2007)
16. Habili, N., Lim, C.C., Moini, A.: Segmentation of the Face and Hands in Sign Language Video Sequences Using Color and Motion Cues. IEEE-TCSVT 14(8), 1086–1097 (2004)
17. Akutsu, A., Tonomura, Y., Hashimoto, H., Ohba, Y.: Video Indexing using Motion Vectors. In: Proc. SPIE Conf. Visual Commu. And Image Processing, vol. 1818, pp. 1522–1530 (1992)
18. Tan, Y.P., Kulkarni, S.R., Ramadge, P.J.: A New Method for Camera Motion Parameter Estimation. In: Proc. ICIP, vol. 1, pp. 406–409 (1995)
19. Kobla, V., Doermann, D., Lin, K.-I.: Archiving, Indexing, and Retrieval of Video in the Compressed Domain. In: Proc. SPIE, vol. 2916, pp. 78–89 (1996)
20. Ewerth, R., Schwalb, M., Tessmann, P., Freisleben, B.: Estimation of Arbitrary Camera Motion in MPEG videos. In: Proc. ICPR., vol. 1, pp. 512–515 (2004)
21. Dante, A., Brookes, M.: Precise Real-time Outlier Removal from Motion Vector Fields for 3D Reconstruction. In: Proc. ICIP, pp. 393–396 (2003)

Parallel Lossless Data Compression: A Particle Dynamic Approach

Dianxun Shuai

East China University of Science and Technology, Shanghai 200237, China
shdx411022@online.sh.cn

Abstract. This paper presents a new generalized particle model (GPM) to generate the prediction coding for lossless data compression. [1]Local rules for particle movement in GPM, parallel algorithm and its implementation structure to generate the desired predictive coding are discussed. The proposed GPM approach has advantages in terms of encoding speed, parallelism, scalability, simplicity, and easy hardware implementation over other sequential lossless compression methods.

1 Introduction

With the ever increasing growth of the Internet and information technology, the data compression (e.g. image compression and text compression) is essential to many areas, such as network communication, multimedia, database, information engineering, image processing and so forth. The data compression is usually categorized into two fundamental types: lossless and lossy. In lossless compression, the reconstructed data after compression is numerically identical to the original data on an entry by entry basis. Whereas in lossy compression, the reconstructed data contains degradations relative to the original. In general, as compared with lossless compression, much higher compression ratio can be achieved by lossy compression. However, the lossless data compression is required or highly desired by many applications where perfect accuracy is needed, such as medical imaging, remote sensing, image archiving, and document reproducing.

The lossless compression strategies based on various transform coding techniques usually result in a negligibly small compression ratio. Therefore, the lossless coding schemes are generally based on relatively simple predictive coding techniques. The state-of-the-art lossless compression techniques such as CALIC [1], BTPC [2], the S+P transform [3-4], the lossless integer wavelet transform [5], and the modified Haar transform [6] are all based on predictive coding techniques. Prediction-based lossless data compression schemes often consist of two distinct and independent parts: modelling and coding. The modelling part is devoted to completely capturing the statistical redundancy of the data source by some kind of model (e.g., prediction model, multi-resolution model, etc.) so as to learn the next

[1] This work was supported by the National Natural Science Foundation of China under Grant No.60575040 and No.60473044.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 266–274, 2008.
© Springer-Verlag Berlin Heidelberg 2008

unknown data entry from the previously observed entries. The coding part is relatively simpler, and often, according to the characteristics of the transformed data, chooses a suitable entropy coding form, e.g., Huffman encoding, arithmetic encoding or Lemp-Ziv encoding. The one-dimensional cellular automata (CA) have been used to generate a predictive coding for lossless compression of a given symbolic string. But the design of local functions for cellular state evolution is very difficult for a lossless compression CA with larger neighbor radius than three, and local functions proper is also very complicated.

This paper is devoted to a new generalized particle model (GPM) to quickly generate the predictive coding for lossless compression of a given symbolic string. GPM transforms the lossless compression process into a parallel motion process of particles.

2 Generalized Particle Model

The generalized particle model (GPM) is discrete dynamic system and is composed of a massive aggregate of particles that are distributed in a one-dimensional array, with each element in the array at most occupied by a particle.

The conceptual architecture of a one-dimensional GPM for the prediction-based lossless data compression is shown in Fig.1. The sampling data are concurrently quantized in a parallel quantizer, and then parallelly sent to a particle array, with each particle carrying a specific data entry and its positional coordinate in the data string. After that, the data string in the particle array is converted into the desired predictive coding through a predictive phase and a transitive phase. Finally, the predictive coding in GPM is output to transmission channel through output port. In the scheme, the particle array perform both as an encoder and as an output buffer.

In general, a predictive coding is composed of two codings: data coding and coordinate coding, that provide all the non-redundant data entries and their corresponding coordinates in the originally input data string, respectively. Accordingly, each particle p_i in a one-dimensional particle array carries two kinds of information: s_i is used for data coding, and v_i is for coordinate coding. The state $\langle\ s_i(t), v_i(t)\ \rangle$ of the particle p_i in GPM is a pair composed of two state

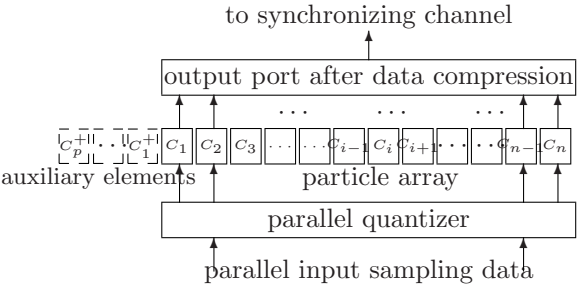


Fig. 1. The Architecture of generalized particle model for prediction-based lossless data compression

components, $s_i(t)$ and $v_i(t)$. Initially, s_i is the i-th data entry of a given input data string, and v_i is set to the value i. The number of possible s_i's is determined by the data type under consideration, e.g., two for binary data, and m for the data defined over the set of m characters or for quantized signal with m quantizing intervals. Whereas the maximal value of v_i depends upon the length of the originally given data string. Every particle can communicate only with its neighbor particles. All the particles in the one-dimensional particle array move synchronously according to the same local function that only involves the neighbor particles. Once GPM evolves to a stable equilibrium point, that is, all the particles don't move any longer, the state set $\{\langle\ s_i, v_i\ \rangle\}$ of particles offers the required predictive coding, where $\{s_i\}$ and $\{v_i\}$ correspond to data coding and coordinate coding, respectively.

To produce a predictive coding for lossless data compression, we design two kinds of local functions: predictive function $\mathcal{P}$ for predictive phase and transitive function $\mathcal{F}$ for transitive phase. During the predictive phase, the particle p_i that is carrying a redundant data entry will be removed from the one-dimension particle array, and thus no particle is located in the positional coordinate i; whereas the particle p_i with a non-redundant data entry will remain its state $\langle\ s_i, v_i\ \rangle$ unchanged. On the other hand, however, during the transitive phase, every particle p_i may concurrently move and thus change the component v_i of the state $\langle\ s_i, v_i\ \rangle$, but doesn't change the component s_i.

The local predictive function $\mathcal{P}$ simultaneously exerts on all the particles in GPM to detect in parallel all the redundant data entries. Suppose that $s_1(t_0), s_2(t_0), \cdots, s_n(t_0)$ are data entries of an initially input data string, and particles , $p_1, p_2, \cdots, p_n$, have initial states, $\langle\ s_1(t_0),1\ \rangle$, $\langle\ s_2(t_0), 2\ \rangle, \cdots, \langle\ s_n(t_0), n\ \rangle$, respectively. Using $s_{i-k}(t_0), \cdots, s_{i-1}(t_0)$, the k-order local predictive function $\mathcal{P}(s_{i-k}(t_0), \cdots, s_{i-1}(t_0))$ predicts the i^{th} data entry so as to obtain $\hat{s}_i(t_0)$, and if the predicted $\hat{s}_i(t_0)$ is equal to $s_i(t_0)$ then it turns out that $s_i(t_0)$ is redundant, and hence the particle p_i should be removed from the one-dimension particle array, in other words, the coordinate i is occupied by an empty particle whose state is denoted by $\langle\ *,\ \emptyset\ \rangle$. We thus have

$$\langle\ s_i(t_0+1), v_i(t_0+1)\ \rangle = \begin{cases} \langle\ *,\ \emptyset\ \rangle & \text{if } s_i(t_0) = \hat{s}_i(t_0), \\ \langle\ s_i(t_0),\ v_i(t_0)\ \rangle & \text{otherwise} \end{cases} \tag{1}$$

where $\hat{s}_i(t_0) = \mathcal{P}(s_{i-k}(t_0), \cdots, s_{i-1}(t_0))$; In order for the local predictive function $\mathcal{P}$ to be able to decide whether the leftmost several data entries of the currently input data string in particle array are redundant or not with respect to previously input data string, we need to know some data entries of the preceding input data string, and for the purpose we prepare several auxiliary elements in a one-dimension particle array in Fig.1. The number k of the auxiliary elements depends on what prediction model is used in lossless compression, e.g., for zero-order polynomial prediction model, one auxiliary element, i.e. $k = 1$, is necessary to determine if the leftmost one data entry of the currently input data string is redundant; for the first-order polynomial prediction model, two auxiliary elements, i.e. $k = 2$, are prepared to judge the redundancy of the

leftmost two data entries of the currently input data string; and for second-order polynomial prediction, $k = 3$ is for the leftmost three data entries, and so on.

After the predictive phase of concurrently executing the local predictive function $\mathcal{P}$ on every particle, all the non-empty particles will simultaneously move under the control of a local transitive function $\mathcal{F}$, until they are not able to move any longer, that is, GPM reaches its fixed point. From final states of non-empty particles, we can obtain the desired predictive coding. Let $c_i(t)$ denote the positional coordinate of particle p_i at time t, and the $q_j(t)$ denote a number of particle that is located at the positional coordinate j at time t. The k-order neighbor of the particle p_i at time t is a coordinate set defined by $\mathcal{N}_k(p_i) = \{(c_i(t) - k), \cdots, (c_i(t) - 1)\}$, where the number k is called a neighbor radius. We define an k-order local transitive function $\mathcal{F}_k$ for a non-empty particle p_i by

$$c_i(t+1) = \mathcal{F}_k(c_i(t)) = c_i(t) - \sum_{j \in \mathcal{N}_k(p_i)} n_j, \tag{2}$$

$$\text{where} \quad n_j = \begin{cases} 1 \text{ if } \ \langle\, s_{q_j(t)}(t), v_{q_j(t)}(t) \rangle = \langle *, \emptyset \rangle; \\ \\ 0 \text{ otherwise.} \end{cases} \tag{3}$$

$\sum_{j \in \mathcal{N}_k(p_i)} n_j$ represents the total number of empty particles that are located in the neighbor of particle p_i. Note that $\max \sum_{j \in \mathcal{N}_k(p_i)} n_j = k$. The k-order local transitive function $\mathcal{F}_k$ implies that the non-empty particle p_i moves left at most k coordinate positions from the current coordinate $c_i(t)$. Moreover, under the local transitive function $\mathcal{F}_k$, we define the state transition of non-empty particle p_i as follows:

$$\langle\, s_i(t+1), v_i(t+1) \,\rangle \Rightarrow \langle\, s_i(t), c_i(t) \,\rangle \tag{4}$$

3 Parallel Algorithm and Its Property

In summary, we give the parallel algorithm GPMA for the predictive coding for lossless data compression as follows.

Parallel Algorithm GPMA

Suppose that a quantized data string has been sent into a one-dimension particle array, with the initial state set $\{\langle\, s_i, i \,\rangle\}$. Given the desired local predictive model and the neighbor radius $k \geq 1$ for the local transitive function.

Costep 1. Concurrently execute the local predictive function $\mathcal{P}$ on every particle p_i;

Costep 2. Concurrently execute the local transitive function $\mathcal{F}_k$ on every non-empty particle p_i;

Costep 3. Concurrently compare the state $\langle s_i(t+1), v_i(t+1) \rangle$ with $\langle s_i(t), v_i(t) \rangle$ for every non-empty particle p_i;

Step 4. If there is $i \in \{1, \cdots, n\}$, such that $\langle s_i(t+1), v_i(t+1) \rangle \neq \langle s_i(t), v_i(t) \rangle$, then go to Costep 2; otherwise finish successfully.

Example 1. Assume that, in the current τ-th input period, a one-dimension particle array is carrying the following τ-th data string

$$S \ = \ s_1 s_2 \cdots s_i \cdots s_{12} s_{13} \ ,$$

with each data entry s_i being a real number. Suppose that using the following first-order polynomial predictive formula

$$\mathcal{P}(s_{i-2}, s_{i-1}) = s_{i-1} + \triangle s_{i-1},$$

where $\triangle s_{i-1} = s_{i-1} - s_{i-2}$, gives rise to nine redundant data entries, $s_2, s_3, s_4, s_6, s_8, s_9, s_{11}, s_{12}, s_{13}$, that is, they are equal to those predicted by the above predictive formula. The auxiliary elements, C_0 and C_{-1}, are used to store the s_{13} and s_{12} of the $(\tau-1)$-th input data string and to obtain the predictive data entries, $\hat{s}_1$ and $\hat{s}_2$ for the τ-th input data string. After executing the local predictive function $\mathcal{P}$ simultaneously on every particle p_i, we have

$$S(1) = \ s_1 * * * s_5 * s_7 * * s_{10} * * * .$$

Then by exerting the local transitive function $\mathcal{F}_2$ four times, the one-dimension particle array respectively has

$$\begin{aligned} S(2) &= \ s_1 * s_5 * * s_7 * s_{10} * * * * *, \\ S(3) &= \ s_1 s_5 * s_7 * * s_{10} * * * * * *, \\ S(4) &= \ s_1 s_5 s_7 * s_{10} * * * * * * * *, \\ S(5) &= \ s_1 s_5 s_7 s_{10} * * * * * * * * * . \end{aligned}$$

It can be seen that the desired data coding $Y = s_1 s_5 s_7 s_{10}$ is obtained from the prefix of $S(5)$. The local function $\mathcal{F}_2$ only needs four time steps to get the Run Length coding. By contrast, the conventional lossless compression method will spend 13 time steps in forming $s_1 s_5 s_7 s_{10}$ in a buffer. Thus, for this example, by using the local function $\mathcal{F}_2$ the GPM approach is more than three times as fast as the conventional compression method.

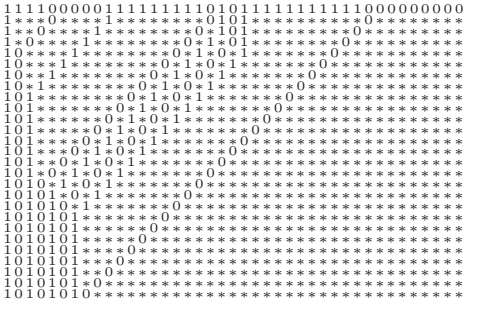


Fig. 2. The evolution to form the data coding 10101010 in a one-dimension particle array under the local functions $\mathcal{P}$ and $\mathcal{F}_1$. Note that the number of the $*$ symbols between two non-redundant symbols, i.e. the number of empty particles, is time-varying, and hence the transition process is entirely different from ordinary homogeneous left-shifting process.

Example 2. Given the currently input binary string

$$S = s_1 \cdots s_{41} = 11110000011111111101011111111111000000000.$$

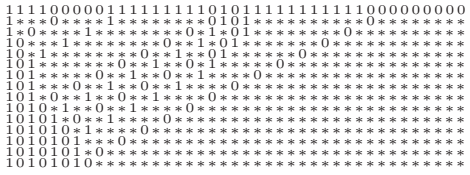


Fig. 3. The evolution to form the data coding 10101010 in a one-dimension particle array under local functions $\mathcal{P}$ and $\mathcal{F}_2$

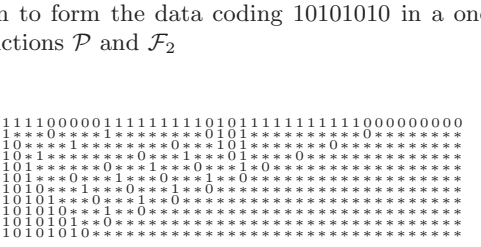


Fig. 4. The evolution to form the data coding 10101010 in a one-dimension particle array under the local functions $\mathcal{P}$ and $\mathcal{F}_3$

Assume that the zero-order polynomial prediction formula $\hat{s}_i = \mathcal{P}(s_{i-1}) = s_{i-1}$ is applied, and $s_1 \neq s_0$, where s_0 is the rightmost entry s_{41} of the preceding input binary string. Since $\hat{s}_1 = \mathcal{P}(s_0) = s_0$, then $s_1 \neq s_0$ implies that $s_1 \neq \hat{s}_1$, and thus s_1 is not redundant. The conventional sequential method will take 41 time-steps for forming the data coding 10101010 and coordinate coding 1 5 10 19 20 21 22 33 in a transmission buffer.

By contrast, the composition of two local functions, $\mathcal{P}$ and $\mathcal{F}$, needs much less time steps to form the same run-length coding in a one-dimension particle array. Executing the predictive function $\mathcal{P}$ concurrently on every particle gives rise to

$$S(1) = 1 * * * 0 * * * * 1 * * * * * * * * 0101 * * * * * * * * * * 0 * * * * * * * *.$$

Then, the first-order $\mathcal{F}_1$, second-order $\mathcal{F}_2$, and third-order $\mathcal{F}_3$ will spend 25, 13 and 9 time-steps, respectively, for GPM to reach the stable fixed-point that corresponds to the data coding 10101010 and coordinate coding 1, 5, 10, 19, 20, 21, 22, 33. Their evolutions to form the above data coding are shown in Fig. 2, 3 and Fig.4, respectively. It turns out that for this example the $\mathcal{F}_1$ and $\mathcal{F}_2$ are 1.6 times and 4.6 times, respectively, as fast as the conventional compression method.

Example 3. In the current τ-th input period, given the following character string in a one-dimension particle array

$$S = s_1 \cdots s_{16} = aaaaaggggghhhbbbb,$$

with each s_i being a character, we use the zero-order polynomial predictive formula $\hat{s}_i = \mathcal{P}(s_{i-1}) = s_{i-1}$. Note that the auxiliary character s_0 is carried by the last particle p_{16} in the $(\tau-1)$-th input period. And assume that $s_0 \neq s_1$, that

is, the character s_1 is not redundant. By executing the local predictive function $\mathcal{P}$ simultaneously on every particle p_i, we have

$$S(1) = a * * * * g * * * h * * b * **,$$

Then by executing the local transitive function $\mathcal{F}_3$ four times, we respectively have

$$S(2) = a * g * * * h * * * b * * * * *,$$
$$S(3) = ag * h * * * b * * * * * * * *,$$
$$S(4) = agh * b * * * * * * * * * * *,$$
$$S(5) = aghb * * * * * * * * * * * *.$$

The desired character coding $Y = aghb$ is obtained from the prefix of $S(5)$. Therefore, the local function $\mathcal{F}_3$ only needs four time steps to form the required coding. Whereas the $\mathcal{F}_2$ and $\mathcal{F}_1$ will only take 6 and 12 time steps, respectively, to get the desired coding $aghb$. By contrast, the conventional method will spend 16 time steps in sequentially forming $aghb$ in a buffer. Therefore, for this example, using $\mathcal{F}_3$ is four times as fast as the conventional lossless compression method.

Example 4. As an extreme case, consider the following the τ-th input character string in a one-dimension particle array

$$S = s_1 \cdots s_m = bbbb \cdots bbbb.$$

The conventional method will spend m time steps in forming the run-length coding b in a transmission buffer. Assume that the above zero-order polynomial predictive formula is used and $s_0 \neq s_1$. After executing the predictive function $\mathcal{P}$ simultaneously on every s_i, we have

$$S(1) = b * * * \cdots * * * .$$

By using the $\mathcal{F}_1$, we have

$$S(2) = b * * * \cdots * * * *.$$

Thus, using the local function $\mathcal{F}_1$ is m times as fast as the conventional compression method.

Remarks 1

• There are two consecutive phases for GPM to form a desired predictive coding. In the predictive phase, i.e. in Costep 1 of the algorithm GPMA, the local predictive function $\mathcal{P}$ concurrently identifies all the redundant data entries carried by a one-dimension particle array, and then concurrently remove their corresponding particles from the particle array. In the transitive phase, i.e. in Costep 2 of the algorithm GPMA, the local transitive function $\mathcal{F}$ is carried out for every non-empty particle, until a fixed-point reaches.

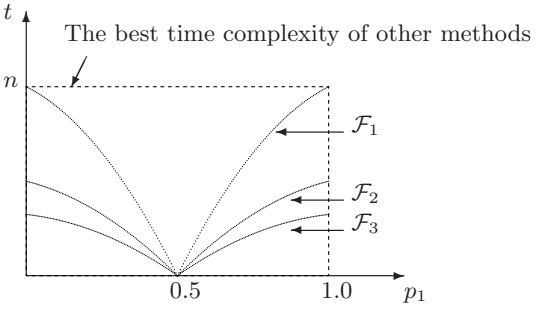


Fig. 5. The time complexity as a function of occurrence probability p_1 of the value 1 in a binary string, where n is the length of a binary string

• From the above examples, we can see that the local function $\mathcal{F}$ will lead to a macroscopically uneven shift behavior: the non-empty particles may move left, which is essentially different from the traditional homogeneous left-shifting operation. Moreover, how to move a non-empty particle depends completely upon the situation of its neighbor, and thus the distinct non-redundant particle may move left unequally at the same time step.

• In principle, various higher-order polynomial prediction formulae may be used for the local predictive function $\mathcal{P}$. Meanwhile, corresponding auxiliary elements in a particle array should be equipped for the state prediction of several leftmost particles.

• The local transitive function $\mathcal{F}$ can find out a predictive coding , at best, n times faster than the conventional lossless compression method, where n is the length of the input data string. The higher the order k of local transitive function $\mathcal{F}_k$, the faster the compression speed. The transitive function described by Eqs.(2),(3),(4) is suitable for arbitrarily large order $k \leq n$.

• Assume that 0 and 1 occur in a binary string at the probability p_0 and p_1, separately. For $\mathcal{F}_\infty, \mathcal{F}_\in$, and $\mathcal{F}_\ni$, the relations between the time complexity of GPMA and the probability p_1 are shown in Fig.5.

In what follows we discuss the correctness and convergency of the algorithm GPMA. Suppose that executing a specified local predictive function $\mathcal{P}$ on every particle of a one-dimension particle array gives rise to

$$S = *^{r_1} w_1 *^{r_2} \cdots w_{m-1} *^{r_m} w_m *^{r_{m+1}} \in Q^n, \tag{5}$$

at time t, where $m + \sum_{i=1}^{(m+1)} r_i = n$; Particles, $w_1, \cdots, w_m$, are non-empty. r_i is the number of empty particles between two non-empty particles, w_{i-1} and w_i, and $*^{r_i}$ represents a consecutive empty-particle string with length r_i. Q is a finite set of symbols over which a data string is defined.

Theorem. The algorithm GPMA will make Eq.(5) converge to a unique fixed-point $S^+ = p_1 p_2 \cdots p_{m-1} p_m *^{n-m}$.

Proof is made by induction, details of proof omitted for page limitation.

4 Conclusions

The GPM architecture, algorithm and local functions are all independent of the problem size, and thus can be easily implemented in VLSI systolic array technology. The proposed GPM approach has the advantages in terms of the higher parallelism, easier VLSI implementation, and the suitability for versatile data and prediction models. In comparison with the cellular automata, the GPM method has much simpler local rules for particle motion than those for cellular state evolution in CA, and moreover the GPM complexity with respect to local rules for particle motion doesn't increase with the neighbor radius.

References

1. Wu, X., Memon, N.: Lossless Interframe Image Compression Via context modelling. IEEE Trans. on Image Processing 9(5), 994–1001 (2000)
2. Robinson, J.A.: Efficient General-Purpose Image Compression with Binary Tree Predictive Coding. IEEE Trans. image Processing 6(4), 601–608 (1997)
3. Said, A., Pearlman, W.: A new fast and efficient image codec Based on Set Partitioning in Hierarchical Trees. IEEE Trans. on Circuits and Systems for Video Technology 6(6), 243–250 (1996)
4. Said, A., Pearlman, W.: An Image Multi-Resolution Representation for Lossless and Lossy Compression. IEEE Trans. on Image Processing 5, 1303–1310 (1996)
5. Calderbank, A.R., Daubechies, I., Sweldens, W., Yeo, B.: Lossless Image Compression using Integer to Integer Wavelet transforms. In: Proc. IEEE Int. Conf. IP, October, vol. 1, pp. 596–599 (1997)
6. Thornton, M.A.: Modified Haar Transform Calculation using Digital Circuit Output Probabilities. In: Proc. of the IEEE Int. Conf. on Information, Communications and Signal Processing, Singapore, September, vol. 1, pp. 52–58 (1997)
7. Wolfram, S.: Random Sequence Generation by Cellular Automata. Advances in Applied Mathematics 7, 123–169 (1986)
8. Zhang, S., Byrne, R., et al.: Quantitative Analysis for Linear Hybrid Linear Cellular Automata and LFSR as Built-in Self-test Generators for Sequential Faults. J. Electronic Testing: Theory and Applications 7(3), 209–221 (1995)
9. Serra, M., Slater, T., et al.: The Analysis of One-Dimensional Linear Cellular Automata and Their Aliasing Properties. IEEE Trans. Computer-Aided Design 9(7), 767–778 (1990)
10. Cattell, K., Zhang, S.: Minimal Cost One-Dimensional Linear Hybrid Cellular Automata of dDegree Through 500. J. Electronic Testing: Theory and Applications 6(2), 255–258 (1995)
11. Cattell, K., Muzio, J.C.: Synthesis of One-Dimensional Linear Hybrid Cellular Automata. IEEE Trans. Computer-Aided Design 15(3), 325–335 (1996)
12. Nandi, S., Vamsi, B., et al.: Cellular Automaton as A BIST Structure for Testing CMOS Circuit. IEE Pro. Computers and Digital Techniques 141(1), 41–47 (1994)

Ball Mill Load Measurement Using Self-adaptive Feature Extraction Method and LS-SVM Model

Gangquan Si, Hui Cao, Yanbin Zhang, and Lixin Jia

School of Electrical Engineering, Xi'an JiaoTong University
Xi'an, Shaanxi, 710049, P.R. China
{gangquan,huicao,lxjia,ybzhang}@mail.xjtu.edu.cn

Abstract. Ball mill load is the most important parameter optimized in the Ball Mill Pulverizing System (BMPS) in Thermal Power Plant. The accurate measurement of ball mill load is imperative and difficult. The approach based on self-adaptive feature extraction algorithm for noise signal and LS-SVM model is proposed to achieve this purpose. By analyzing the sensitivity distributions of working condition transitions, the characteristic power spectrum (CPS) is obtained. Then the self-adaptive weights can be calculated based on the CPS and the centroid frequency. The extracted features of the mill noise using this method show good adaptability to working condition transitions. Further more, softsensing models are built by combining a LS-SVM model and various feature extraction methods so as to estimate the mill load. Experimental results show that the performance of the model combined with the proposed extraction method is better than that with other methods.

Keywords: Ball mill pulverizing system, self-adaptive feature extraction, least squares support vector machines, softsensing.

1 Introduction

Mill load is an important operating parameter in ball mill pulverizing system(BMPS), which can represent the running conditions. By its optimization, it is possible to produce significant improvements in production capacity and energy efficiency and to maintain the BMPS under optimal conditions [1]. Mill Load measurement at laboratory scale may be carried out by a weighing platform easily. However, it is infeasible at industrial scale and the measurement demands more sophisticated measuring techniques [2]. Many measurement methods have been presented in the past decades. Power consumption of the mill motor is first used as mill load to optimize the milling efficiency [3]. The sensitivity and accuracy are too poor to be used successfully because that the power consumptions of various filling states of the mill are similar. The operators usually use the inlet-outlet difference pressure to control the ball mill. However, the inlet-outlet difference pressure measurement is often affected by the variations of ventilation rate. For tracking particle motion, a novel method by using diagnostic X-rays from a bi-planar angioscope is used in a laboratory mill [4]. It is an accurate

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 275–284, 2008.
© Springer-Verlag Berlin Heidelberg 2008

technique but the investment cost at industrial scale is large. Softsensing technique is a newly and alternatively approach to solve these problems. It estimates the primary variable by combining the secondary variables, which are low cost and easy measurement.

One of the principal variables for estimating the mill load is the noise signal [5,6]. The noise signal processing is rather complicated and the features of the signal must be extracted exactly before using in the softsensing model. There are mainly two methods to process the noise signal. One method is using the energy of all frequency components of the noise signal as the features[5,7](named as EWAPS method). The other considers that there exists a continuous feature frequency band (FFB) and only those frequency components in FFB are effective[6](named as EWFBB method). Then the features can be calculated by summating the energy of the frequency components in FFB. Recently, a lot of experiments are implemented and large amounts of noise data from various mills are gathered. After a further analysis, we consider that the FFB is discontinuous and the sensitivities of the effective frequency components are also different. Further more, as working conditions of the pulverizing mill transited, the parameters of the noise signals are also changed. Based on these analyses, we proposed a self -adaptive feature extraction method for noise signal.

The key of the softsensing technique is how to design an efficiency model, which can estimate the target function accurately. At the same time, the model should have low complexity and good generalization performance. A lot of novel methods have been used to create the softsensing model, such as neural networks [8], adaptive neural-network based fuzzy inference system [9], partial least squares [10], etc. All the approaches mentioned above are almost based on the empirical risk minimization principle. These algorithms often lead to the problem of over fitting. Based on statistical learning and the structural risk minimization principle, the support vector machines (SVM) is presented in [11], which considers both empirical risk and generalization performance, and it can be used to approximate nonlinear functions. The LS-SVM method whose objective function includes an additional squared error sum term is presented in [12]. Many successful applications [13,14] demonstrate that LS-SVM is especially suitable for softsensing model.

In this paper, a novel self-adaptive feature extraction method based on CPS and centroid frequency is proposed and the LS-SVM method with a radial basis function kernel is employed to build a mill load softsensing model. By analyzing the sensitivity distributions of working condition transitions, the ineffective frequency components are removed and the CPS is obtained. The self-adaptive weights are calculated based on the centroid frequency and the distribution of the CPS. Then multiplying the energy of the CPS by the self-adaptive weights, the features of mill noise can be extracted. Finally, a sparse LS-SVM based model is used to estimate the mill load and achieves a good effect.

The rest of this paper is organized as follows. Section 2 explains the self-adaptive feature extraction method in detail. Section 3 reviews the LS-SVM

modeling method. Section 4 gives practical experiments to demonstrate the effectiveness of the proposed method. The conclusions are given in section 5.

2 Self-adaptive Feature Extration

Precise feature extraction for the noise signal of the ball mill is important in improving the estimated precision and the adaptability of the mill load softsensing model. Therefore, we proposed a self-adaptive feature extraction method based on the centroid frequency of the CPS. The details of the self-adaptive feature extraction method are described below.

2.1 Characteristic Power Spectrum

The noise signal of ball mill consists of the inherent noise of the mechanical devices, the impact noise, the motor noise, the gear noise, the background noise, etc [15]. In all the noise sources, the impact noise which is generated by the impacts of ball to ball, ball to coal, and ball to lining is effective. The other noise is independent of the mill load and must be removed.

Suppose that a data sequence of noise signal under the working condition $state(i)$

$$\{x_{state(i)}(n); n = 1, 2 \cdots N, i = 1, 2 \cdots I\} \tag{1}$$

where $x_{state(i)}(n)$ is the nth sample under the working condition $state(i)$, N is the sample number, and I is the number of working conditions. Then the CPS can be obtained as follows:

Step 1: Calculating the power spectrum of the noise data sequence under working condition $state(i)$

$$P_{state(i)}(k) = |X_{state(i)}(k)|^2/N \tag{2}$$

where $\{X_{state(i)}(k), k = 1, 2 \cdots N\}$ is the discrete Fourier transform of $\{x_{state(i)}(n)\}$, computed with a N-point fast Fourier transform algorithm.

Step 2: Calculating the sensitivity distribution of working condition transitions from $state(i)$ to $state(j)$

$$s_{state(i \to j)}(k) = \left| \frac{P_{state(i)}(k) - P_{state(j)}(k)}{max\{P_{state(i)}(k), P_{state(j)}(k)\}} \right| \tag{3}$$

clearly, larger $s_{state(i \to j)}$ signifies that the corresponding frequency component is better to estimate the mill load.

Step 3: Selecting the sensitivity threshold and picking out the feature frequency components under the same working condition transitions. Defining the set of feature frequency which meeting the sensitivity constraint as follows.

$$K_f(state(i \to j)) = \{k, s_{state(i \to j)}(k) > \lambda_{state(i \to j)}\} \tag{4}$$

subject to the constraints

$$k = 0 - k_h, k_h = min\{[f_h/\Delta f], [f_s/2\Delta f]\} \tag{5}$$

where $[\cdot]$ denotes round operation. $\lambda_{state(i\to j)}$ is the sensitivity threshold under the working condition transitions from $state(i)$ to $state(j)$. f_h is the high limit frequency of the actual signal, which is determined by the sensor. f_s is the sampling frequency, which is determined by both the data acquisition module and the software. Δf is the frequency resolution and which is calculated by $\Delta f = f_s/N$.

Step 4: obtaining the set of feature frequency under all working condition transitions as follows.

$$K_f = K_f(state(1 \to 2)) \cap \cdots \cap K_f(state(i \to j)) \tag{6}$$

Thus, the most superior frequency components can be selected according to the sensitivity distributions.

Step 5: Finally obtaining the CPS

$$P_f(k) = \{P(k), k \in K_f\} \tag{7}$$

2.2 Self-adaptive Feature Extraction

After obtaining the CPS, the effective frequency components which represent the impact noise have been identified. However, the sensitivity of different frequency component is unequal. The frequency component with high-sensitivity should be more important and have bigger weight in the feature energy calculation. It is found that the frequency component with high-sensitivity usually near the centroid frequency by analyzing large amount of various mill noise signals. Thus we propose a self-adaptive weight based algorithm as follows:

Step 1: Calculating the centroid frequency of the CPS

$$k_c = (\sum kP_f(k))/(\sum P_f(k)) \tag{8}$$

Step 2: Calculating the self-adaptive weights. We select Gaussian combination function as the weight function.

$$\mu(k) = \begin{cases} exp(-(k-c_1)^2/2\delta_1^2) & k < c_1 \\ 1 & c_1 \le k \le c_2 \\ exp(-(k-c_2)^2/2\delta_2^2) & k > c_2 \end{cases} \tag{9}$$

Step 3: Calculating the parameters of the weight function

$$c_1 : \max_{k_{c1}}((\sum_{k=k_{c1}}^{k_c} P_f(k))/(\sum_{k=0}^{k_c} P_f(k)) \ge \alpha_1) \tag{10}$$

$$\delta_1 : \min_{\Delta k_1}((\sum_{k=c_1-\Delta k_1}^{k_c} P_f(k))/(\sum_{k=0}^{k_c} P_f(k)) \ge \alpha_2) \tag{11}$$

$$c_2 : \max_{k_{c2}}((\sum_{k=k_c}^{k_{c2}} P_f(k))/(\sum_{k=k_c}^{k_h} P_f(k)) \ge \alpha_3) \tag{12}$$

$$\delta_2 : \min_{\Delta k_2}((\sum_{k=k_c}^{c_2+\Delta k_2} P_f(k))/(\sum_{k=k_c}^{k_h} P_f(k)) \geq \alpha_4) \tag{13}$$

where parameters $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ are the threshold, which are determined according to the actual experiment data and experiences. The centroid frequency k_c divides the CPS into two parts. $\{c_1, \delta_1\}$describes the energy distribution of the left side of the CPS and $\{c_2, \delta_2\}$ the right side.

Step 4: Calculating the feature energy of the noise signal

$$E = \sum \mu(k) P_f(k) \tag{14}$$

Now, the features of the noise signal are extracted, which can be used in the softsensing model to predict the mill load. The effect of the proposed feature extraction method is shown in section 4.

3 LS-SVM Based Softsensing Model

LS-SVM is presented on the basis of classical SVM and it tries to minimize primal cost function subject to equality constraints instead of inequality ones. Therefore, LS-SVM solves a set of linear equations instead of computational cost quadratic programming problem and has a better efficiency than standard method. In the following, a LS-SVM based softsensing model for mill load measurement is constructed. More details about LS-SVM can be referred to reference [16].

Assume a training set is given like $\{x_k, y_k\}_{k=1}^N$, with the input $x(k) \in R^n$ and the output $y(k) \in R$. The regression model $y(x) = w^T \varphi(x) + b$, as introduced in reference [16], can be obtained by solving the following optimization problem

$$\min_{w,b,e} J(w, e) = \frac{1}{2} \parallel w \parallel_2 + \frac{C}{2} \sum_{k=1}^{N} e_k^2 \tag{15}$$

subject to the constraints

$$y_k = w^T \varphi(x) + b + e_k, k = 1, 2 \cdots N \tag{16}$$

where C is the user-defined regularization constant which balance the model's complexity and approximation accuracy, and e_k the training error. $\varphi(\cdot)$ maps the input data to a higher dimensional feature space.

Using Lagrange function $L(w, b, e, \alpha)$ and radial basis function kernel $K(x, x_k) = exp\{-\|x - x_k\|^2/\delta^2\}$,the LS-SVM regression model can be obtained as

$$y(x) = \sum_{k=1}^{N} \alpha_k K(x, x_k) + b \tag{17}$$

Combining the input variables of the power consumption method, the inlet-outlet difference pressure method and the outlet temperature method, the following input output model

$$y_{ml} = f(x_{en}, x_{mi}, x_{dp}, x_{ot}) \tag{18}$$

is constructed to measure the mill load of the ball mill. y_{ml} denotes the mill load. x_{en} denotes the features extracted from the noise signal. x_{mi} ,x_{dp} and x_{ot} denote the motor current, the outlet temperature and the inlet-outlet difference pressure respectively. $f(\cdot)$is a nonlinear function which are realized by the LS-SVM method.

In the softsensing model based on LS-SVM, in order to get the sparse support vector set, we adopt the procedure proposed by Suykens [17]. In the procedure, the unimportant support vectors which have small α are removed. So the complexity of the model is reduced and the generalization performance is improved. We also used the Leave One Out Cross-Validation method to find out the optimal parameter values of the model. All the aforementioned calculations were performed using MATLAB 7.04 and the free LS-SVM toolbox V1.5 [16].

4 Experiments and Results

The self-adaptive feature extraction method and the LS-SVM based softsensing model mentioned above have been applied to the BMPS of Qinling Power Plant in China. The type of the ball mill is DTM350/700, whose diameter is 3.5 meters and length 7.0 meters. A lot of experiments are implemented to collect the data used in this work and the results are described below.

4.1 Noise Signal Processing

The noise data of low-level filling state, middle-level filling state and high-level filling state are obtained, and the corresponding working conditions are named as state I, state II and state III respectively. The sensitivity distributions are showed in Fig.1. In the procedure of calculating the CPS, $\lambda_{state(I \rightarrow II)} = 0.20$ and $\lambda_{state(II \rightarrow III)} = 0.18$ are selected by trial way. The sensitivity threshold line and the high limit frequency line partition the figure into four zones, and the frequency components which are in left upper zone compose the feature set. The overlap of the two feature sets forms the frequency variable of the CPS. Fig .2 shows the CPS, the centroid frequency and the weights of the working conditions state I, state II and state III. And the parameters are presented in Table 1. The centroid frequency is decreasing as the mill load increasing, which is 4,000Hz in state I and 3,350Hz in state III.

To illustrate the performance of our proposed feature extraction method (SAWCPS method), three other feature extraction methods are adopted to

Table 1. Parameters of self-adaptive feature extraction algorithm

Working conditions	k_c	c_1	δ_1	c_2	δ_2
StateI	80	62	31	102	41
StateII	73	58	38	86	56
StateIII	67	54	40	77	60

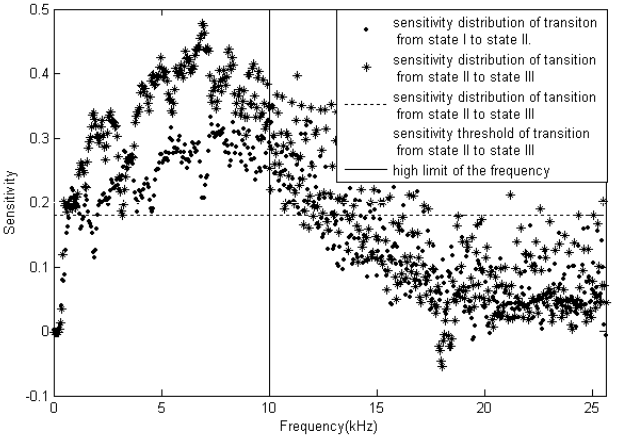


Fig. 1. The sensitivity distribution

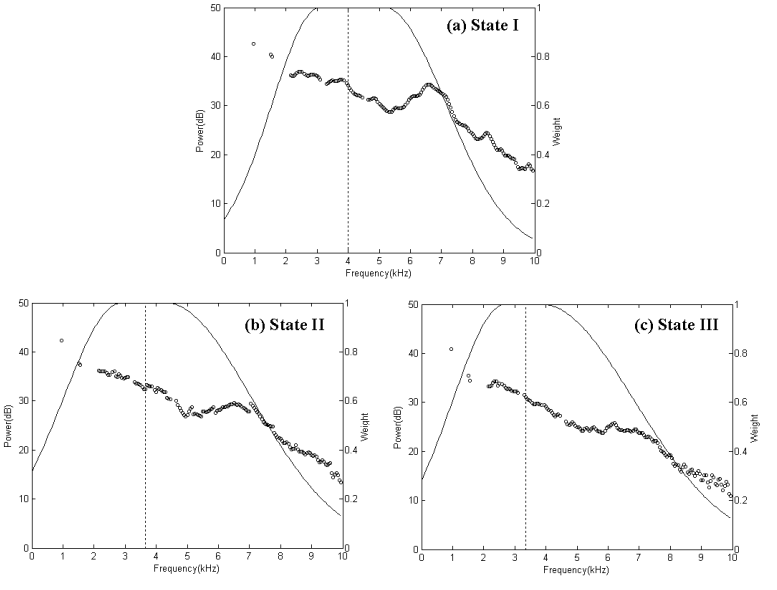


Fig. 2. The CPS, the centroid frequency and the weights of different working conditions.Cycle "o"denotes the CPS. Solid line "-"denotes the weights. Dashed line "···"denotes the centroid frequency.

the same noise signal. The calculation equations of these three methods are: (1) EWAPS method: $E_{EWAPS} = \sum P(k), k = 1 \cdots k_h$; (2) EWFFB method: $E_{EWFBB} = \sum P(k), k = k_{FBl} \cdots k_{FBh}$.where k_{FBl} and k_{FBh} denote the low

boundary and high boundary of the FBB; (3) EWCPS method: $E_{EWCPS} = \sum P_f(k), k = k_1 \cdots k_h$.

The noise data are collected in the process of mill load linearly increasing. The feature results are presented in Fig. 3. For comparison simply, the feature values of the four methods are normalized. It is known that the features of noise should decrease linearly as mill load increasing linearly. As shown in Fig.3, the features of EWAPS method are almost no changes and can not be used to estimate the mill load. The reason of this phenomenon is that the energy of ineffective frequency components is larger than the effective ones and the energy of ineffective frequency components are almost not change as the mill load increasing. The features of EWFBB method decrease as the mill load increase. However, the features have large disturbance which may be generated by the ineffective frequency components in FFB. We can draw that from Fig.1. The features of EWCPS method meet our expectation better but the linearity is poor. The features of our proposed SAWCPS method match the process of the experiment preferably, which has the advantages of low disturbance, high sensitivity and good linearity.

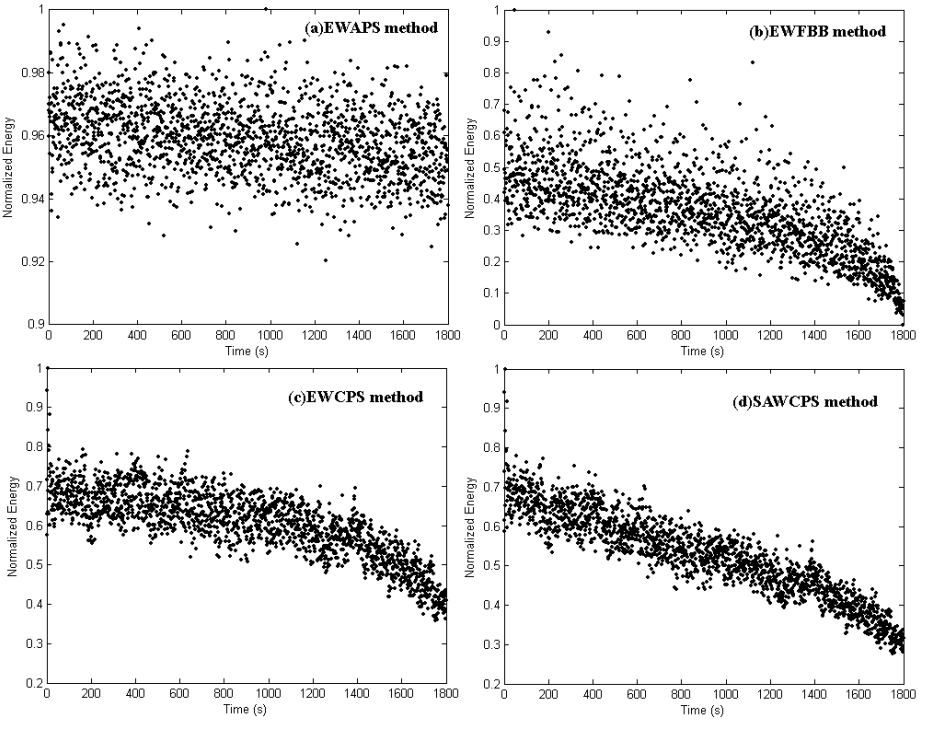


Fig. 3. Results of features extracted from mill noise by using different feature extraction method

4.2 Mill Load Measurement

Total 4800 samples were collected in the experiment, in which the processes of increasing feeding rate, decreasing feeding rate, changing the air draft capacity, etc. are included. 192 samples are selected as the training set and 480 samples the testing set. After removing the least relevant support vectors, there are 114 support vectors remained. The optimal parameters pair of $\{C, \delta_2\}$ was found at the value of $C = 53.039$ and $\delta_2 = 1.3947$.

Four models were developed using LS-SVM and the results were compared. The differences among the four models are the feature extraction method of the noise signal. Table 2 compares the RMSE of training data set and the testing data set. It is observed that the RMSE values of both training data sets and testing data sets decrease by using our proposed self-adaptive feature extraction method. By combining the self-adaptive feature extraction method and the LS-SVM model, the measurement of mill load is realized. At the same time, the model achieved a good estimate precision with a good generalization performance.

Table 2. RMSE results of training and testing data

Model	RMSE of training data set	RMSE of testing data set
EWAPS	0.8351	3.2199
EWFFB	0.8034	2.6129
EWCPS	0.6833	1.3518
SAWCPS	0.6267	1.3056

5 Conclusion

The paper studied the ball mill load measurement by combining the self-adaptive feature extraction method and the LS-SVM based softsensing model. By removing the ineffective frequency components from the mill noise signal and calculating the corresponding weight function, the feature values are obtained. The proposed feature extraction method can adapt the working condition transitions and have better correlation with mill load. The LS-SVM softsensing model produced acceptable precision and accuracy in estimating the ball mill load and its effectiveness has been demonstrated in the experiment. Experimental results indicate that the proposed approach have good generalization performance. Finally, it can be concluded that the self-adaptive feature extraction method and LS-SVM model are promising alternative for the measurement of the ball mill load.

Acknowledgment

This work is supported by the National High Technology Research and Development Program of China (No.2006AA04Z180) and the Ph.D. Programs Foundation of Ministry of Education of China (No.20070698059).

References

1. Kolacz, J.: Improving Fine Grinding Process in Air Swept Ball Mill Circuits.Ph.D. Dissertation, University of Trondheim NTH, Department of Geology and Mineral Resources Engineering, Norway, NTH Trykk (1995)
2. Kolacz, J.: Measurement System of the Mill Charge in Grinding Ball Mill Circuits. Minerals Engineering 10(12), 1329–1338 (1997)
3. Olsen, T.O.: Automatic Control of Continuous Autogenous Grinding. In: Conference on Automation in Mining, Mineral and Metal Processing, Johannesburg (1976)
4. Powell, M.S., Nurick, G.N.: A Study of Charge Motion in Rotary Mills. Part II-Experimental Work. Minerals Engineering 9(3), 343–350 (1996)
5. Cheng, Q.M., et al.: The Overview on the Development of Control Techniques on Intermediate Storage Bunker Ball Mill Pulverizing System of Power Plant. Journal of Shanghai University of Electric Power 22(1), 48–54 (2006)
6. Si, G.Q., Cao, H.: Soft Sensing Model Based on ANFIS for Measuring Ball Mill Load of Power Plant. Chinese Journal of Scientific Instrument 28, supp. 2, 1–5 (2007)
7. Zeng, Y., Forssberg, E.: Monitoring Operating State by Measuring Vibration and Acoustic Signals in An Industraial Scale Ball Mill. In: Conference on Mineral Processing, Lulea, Sweden, pp. 135–146 (1993)
8. Joodaki, M., Kompa, G.: Application of Neural Networks for Extraction of Distance and Reflectance in Pulsed Laser Radar. Measurement 40(6), 724–736 (2007)
9. Ying, L.C., Pan, M.C.: Using Adaptive Network Based Fuzzy Inference System to Forecast Regional Electricity Loads. Energy Conversion and Management 49(2), 205–211 (2008)
10. Min, K.G., Han, I., Han, C.: Iterative Error-based Nonlinear PLS Method of Nonlinear Chemical Process Modeling. Journal of Chemical Engineering of Japan 35(7), 613–625 (2002)
11. Vapnik, V.: Statistical Learning Theory. John Wiley, New York (1998)
12. Suykens, J.A.K., Vandewalle, J.: Least Squares Support Vector Machine Classifiers. Neural Process Lett. 9(3), 293–300 (1999)
13. Salgado, D.R., Alonso, F.J.: An Approach Based on Current and Sound Signals for In-process Tool Wear Monitoring. International Journal of Machine Tools and Manufacture 47(14), 2140–2152 (2007)
14. Li, X., Cao, G.Y., Zhu, X.J.: Modeling and Control of PEMFC Bbased on Least Squares Support Vector Machines. Energy Conversion and Management 47(7-8), 7–8 (2006)
15. Chen, J.: The Noise Control Technique of Ball Mill. China Electric Power Press, Beijing (2002)
16. Suykens, J.A.K., Gestel, T., De Brabanter, J., De Moor, B.: Least Squares Support Vector Machines. World Scientific, Singapore (2002)
17. Suykens, J.A.K., et al.: Weighted Least Squares Support Vector Machines: Robustness and Sparse Approximation. Neurocomputing 2(48), 85–105 (2002)

Object Recognition with Task Relevant Combined Local Features

Wenjun Zhu and Liqing Zhang

Department of Computer Science and Engineering
Shanghai Jiao Tong University, Shanghai 200240, China
{wilsonzhu,lqzhang}@sjtu.edu.cn

Abstract. A number of cortex-like hierarchical models of object recognition have been proposed these years. In this paper, we improve them by introducing supervision during forming combined local features. The traditional cortex-like hierarchical models always contain three layers which imitate the functions of neurons in ventral visual stream of primates. The bottom layer detects orientation information in a local area. Then the middle layer combines these information to form combined features. Finally, the top layer integrates combined features to form global features which are input into a classifier. In these models, three stages to form global features are all unsupervised. The supervision procedure only occurs after global features are generated, which is implemented by the classifier. But we think the supervision should occurs earlier. For a particular object recognition task, the second stage of generating global features is also supervised because only task relevant combinations are useful. In our paper, we analyze why introducing supervision in this stage is necessary. And we explain task relevant combined local features can be extracted by some feature selection algorithms. We also apply this improved system to a series of object classification problems and compare it with traditional models. The simulation results show that our improvement really boosts object recognition performance.

Keywords: object recognition, task relevant, visual cortex, Gabor filter.

1 Introduction

Human brain can recognize what it concerns in clutter within a very short time. This ability surpasses all artificial systems. So understanding how visual cortex recognizes objects and building brain-like recognition system attract many researchers in the fields of physiology, neuroscience and computer science.

Over the last decade, a number of basic properties about object recognition in cortex have been found through many physiological experiments [1]. Visual signals received by retina are processed stage by stage in the ventral visual pathway, which runs from primary visual cortex (V1), over extrastriate visual areas V2 and V4 to inferotemporal cortex (IT). The following properties have been widely accepted in this stearm. Firstly, simple cells in V1 respond perferably to oriented bars [2]. Secondly, neurons along the ventral stream show an increase in receptive field size as well as in the complexity of their preferred stimuli. V4

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 285–292, 2008.
© Springer-Verlag Berlin Heidelberg 2008

provides information about the individual contour elements and these elements are integrated into shapes at the level of area IT [2,3,4,5]. Finally , at the top of the ventral stream, in anterior inferotemporal cortex (AIT), cells are tuned to very complex stimuli such as faces, animals, scenes, etc [6].

Based on these widely accepted properties, several transformation invariant object recognition models have been proposed [7,8,9]. HMAX [10,11] proposed by MIT-CBCL is one of the most significant models among them. This hierarchical model based on feedforward connections ignores any dynamics of the back-projection, although feedback connections really exist in our visual cortex. In the simplest form of the model, it contains fours layers, which are S1,C1,S2 and C2 from bottom to top. S1 detects orientation in a particular position. C1 takes the max over a local area to form local invariant features. S2 combines features output by C1. This procedure is implemented by extracting some patches or templates from positive training images and calculating similarity between the samples and the templates. C2 deals with global invariance by getting the maximal value over the outputs of S2. Finally, the global invariant features formed in C2 are input into a classifier (SVM, neuron network, etc.). The whole recognition procedure is composed of two stages. The stage forming C2 features is unsupervised and the stage using classifier to recognize objects is supervised. However, we find for a particular object recognition task not all C2 features extracted by this unsupervised method are useful. Some of them are task irrelevant. These irrelevant features may reduce the system's performance. So we introduce supervision to form C2 features in our improved model. Task relevant combinations are extracted in the layer S2, because we think only partial combined local features are useful for a particular recognition task. For example, a feature which can discriminate two different people in face recognition task doesn't work in the task to detect whether the object is a face, and vice versa. This improvement boosts the recognition performance, which we will show in Section 3. Another dedication of this improvement is that it avoids the problem in combining features, which is the number of possible combinations is too large. The authors of HMAX resolve this problem by randomly selecting a few ones from all combinations. In our model, we select task relevant combinations. Obviously it is better than random selection.

This paper is organized as follows. In Section 2, we briefly introduce HMAX at first. Then present our improvement and describe detailed implementation of the whole system. In Section 3, we apply our system to both binary and multi-class classification problem on public database and compare the results with HMAX. Finally, we summarize our work and propose some open questions about the limitations and possible improvements of the model in Section 4.

2 Hierarchical Model and Feature Selection

HMAX is one of the most significant hierarchical models. It was first proposed by Riesenhuber and Poggio in [10]. After that, Serre and his co-workers evolved the model and applied it to a number of challenging recognition tasks [11].

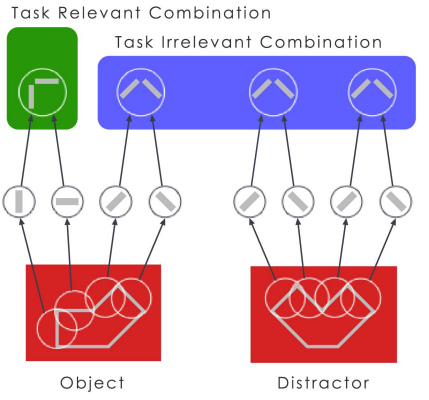


Fig. 1. An example of task relevant and irrelevant combination

HMAX model consists of simple S units and complex C units alternately. In the simplest form, there are four layers (S1, C1, S2, C2). The readers can refer [10,11] for the detailed implementation of these layers.

The difference between our model and HMAX is we employ supervision in the stage of forming combined features. In fact, learning occurs at all stages in ventral visual stream. It is unsupervised in the lower layers and supervised in the higher layers. The properties of simple and complex cells in V1 is learned from natural images by unsupervised proedure [12,13]. In the next stage forming local combined features, the authors of HMAX think it is also unsupervised. However we don't agree this opinion. Maybe this stage seems unsupervised for all kinds of recognition tasks. But for a special recognition task, we believe only a few task relevant combinations are useful and the learning is supervised. In other words, we think there are many cells in cortex corresponding all possible combinations, but only task relevant cells are activated for a particular recognition task and the others can be ignored. We illustrate a example in Fig. 1. The combination in blue rectangle is useless for the object recognition task, and the one in green rectangle is task relevant. In our model, we only keep these task relevant combinations and the others are discarded.

How to extract task relevant features is a feature selection problem. Now, there are many methods to resolve feature selection problem [14,15,16,17,18]. In this version of our model, we use a greedy algorithm to resolve it. The algorithm can be described as:

1. Set $X \leftarrow$ " initial set of n features", $S \leftarrow$ " empty set".
2. $\forall x \in X$, compute $J(\{x\} \cup S)$[1].
3. Find the feature that maximizes $J(\{x\} \cup S)$, set $X \leftarrow F \backslash \{x\}$ and $S \leftarrow \{x\} \cup S$.
4. Repeat until desired number of features are selected.

[1] $J(\bullet)$ is a criterion to evaluate the selected subset of features.

In our experiments, the inter-intra criterion is used, i.e. $J = trace(S_w^{-1} S_b)$, where S_b is between-scatter and S_w is within-scatter.

At the end of this subsection, we summarize our method completely[2]:

1. For every image, we apply a battery of Gabor filters to it. The filters are in 4 orientations θ and 16 scales s. Then we obtain $16 \times 4 = 64$ maps $(S1)_\theta^s$ arranged in 8 bands (e.g., band 1 contains filter outputs of size 7 and 9, in all four orientations, band 2 contains filters outputs of size 11 and 13, etc).
2. Take the max over scales and positions within each band: each band member is sub-sampled by taking the max over a grid with cells of size N^Σ first and the max between the two scale members second, e.g., for band 1, a spatial max is taken over an 8×8 grid first and then across the two scales (size 7 and 9). Note that we don't take a max over different orientations. Hence, each band $(C1)^\Sigma$ contains 4 maps.
3. Extract all possible patches of size $n_i \times n_i \times 4$ from all $(C1)^\Sigma$ of all positive training images. They are $P_{i=1,2,\cdots,K}$.
4. For each C1 image $(C1)^\Sigma$ and each P_i, compute: $Y = \exp(-\gamma \parallel X - P_i \parallel)^2$. X is a patch of $(C1)^\Sigma$, whose size is same as P_i. And we let X run over all positions. Obtain S2 map $(S2)_i^{\Sigma 3}$.
5. Take the max over all positions and bands, we obtain $(C2)_i$. Now, every input image are converted into K C2 features. Every C2 feature correspond a patch P_i in step 3.
6. In the training set, we use our greedy feature selection algorithm to extract k task relevant C2 features ($k << K$). The patches corresponding these features are task relevant combinations.
7. Input these selected C2 features into a classifier (linear SVM in our implementaion).

3 Experimental Results

We evaluate our system by several object detection tasks and a multi-class object classification task in this section. The classifier used in our system is a linear SVM.

The database used in our experiments is Caltech6[4]. This is a public multi-class object database. The Caltech6 database contains six object categories and a background category (used as the negative set). In fact, there are only 5 different kinds of objects in Caltech6 database, because the first two categories are both cars (rear). For cars (rear), we only use the images under the directory *cars_brad* in our experiments.

[2] In this version, the implementation of S1 and C1 are same as HMAX. So you can refer [11] for more information.

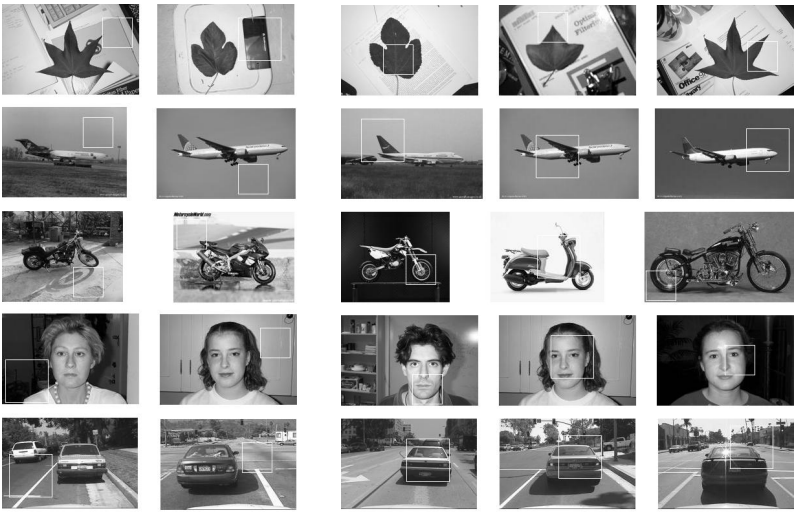
[3] For a train image, we should calculate all S2 features. However we need not compute all S2 features for a test image, because only task relevant patches are useful for them.

[4] http://www.vision.caltech.edu/Image_Datasets/Caltech6/

Fig. 2 illustrates some examples in Caltech6. These images contain the target object embedded in a large amount of clutter. The challenge is to learn from unsegmented images and discover the target object class automatically. As a preprocessing procedure, we normalize all images to 140 pixels in height (width is rescaled accordingly so that the image aspect ratio was preserved) and converted to gray values.

3.1 Object Detection

In our object detection experiments, we selected 40 images in a category as the positive train samples and 50 background images as the negative ones. We also extracted 100 positive and 100 negative test samples from the remaining images. The performance of HMAX was averaged over 10 runs. The code of HMAX in our experiments was downloaded from MIT-CBCL (http://cbcl.mit.edu/software-datasets/). Only 6×6 and 10×10 patches of band 2 were extracted in our experiments for saving computation time (Using patches of more sizes and bands will improve the performance.). In HMAX, we randomly extract 100 patches in each size and totally 200 C2 features. In our system, we calculated correlation coefficient between every C2 features and train label. Then run our greedy feature selection algorithm on the most correlated C2 features (1000 features) to extract task relevant patches (50 patches). Some features extracted by HMAX are task



(a) task irrelevant patches (b) task relevant patches

Fig. 2. (a) and (b) are some examples of task irrelevant and relevant patches. These patches are extracted from C1 outputs. We use white rectangles to denote the areas corresponding these patches in the original images. Task irrelevant patches in (a) come from background or other objects and task relevant patches in (b) come from the target objects.

Table 1. Performance of HMAX and our model on five object detection tasks

Categories	HMAX (200)	Our Model (50)
Leaves	94.9	97.5
Airplanes	91.6	94
Motorbikes	95.8	98
Faces	98.2	99
Cars	97.9	99.5

irrelevant and useless for classification. But in our system almost all features are task relevant. So our system need fewer features than HMAX for the same classification problem. You can see the performance of our system with only 50 features even better than HMAX with 200 features in Table 1.

Fig. 2(a) illustrates some examples of task irrelevant patches extracted by HMAX. We can see these patches come from background or other objects rather than the target object. So the features formed from these patches are useless for recognize the target object. Fig. 2(b) illustrates some patches extracted by our system. We can see these areas are the most important parts of the target objects and features formed from these areas are very discriminative between the objects and distractors. Although size and position of target objects vary in different images, our system can detect where they are and extract features from these places.

Table 1 summarize the experimental results. We use 200 features in HMAX and 50 features in our model. The classification rates of our model are better than HMAX in these challenging applications even with lesser features. And the results show that our system is very efficient in object detection.

3.2 Multi-class Object Classification

In the multi-class object classification, we still used 5 object categories in Caltech6. We used all images of these five categories in our experiment. In each category, we randomly selected 30 samples for training and the rest for testing. As in object detection, we selected the most correlated C2 features (500 features) for each category and extracted task relevant ones (50 features) among them by our feather selection algorithm. We extracted totally 250 patches, 50 patches for each category. During selecting object relevant patches, we regarded the images of this category as positive samples and the others as negative samples.

Table 2. Performance of multi-class object recognition (see text)

Categories	Train	Test	One VS Rest	All
Leaves	30	156	99.6	
Airplanes	30	1044	91.7	
Motorbikes	30	796	96.8	95.0
Faces	30	420	98.8	
Cars	30	496	98.4	

Table 2 is the results of our multi-class object classification experiment. The second and third columns are number of train and test samples respectively. The fourth column is the classification rates between one category and the rest (only 50 features relevant to this category are used). The last column is the performance of our multi-class object recognition task by total 250 features. The results show that our system performs well even with a small number of training samples.

4 Summary and Discussion

In this paper, we have proposed a new object recognition model by introducing supervision to the local feature combination phase of traditional HMAX. HMAX model proposed by MIT-CBCL is a famous hierarchical object recognition model in cortex. But randomly selecting patches in S2 layer is one limitation of it. We find introducing supervision and extracting task relevant patches can resolve the problem effectively. By this improvement, higher recognition performance with less features are achieved (see Section 3).

One limitation of our system is long computation time in training, because computing C2 features for a patch is time consuming and we compute C2 features for all patches on training data. Fortunately, the testing procedure is mush faster than training, for only a few task relevant C2 features need to be calculated (50 features in our experiments).

In HMAX and our model, only shift and scale transformations are considered. In real world, we will face all kinds of other transformations, such as rotation, illumination, occlusion, etc. How to extract invariant features under these complex transformations is a challenging work. In the next step, we will modify our model and try to resolve these more challenging problems.

Acknowledgments. The work was supported by the National High-Tech Research Program of China (Grant No. 2006AA01Z125) and the national natural scinece foundation of China (Grant No. 60775007).

References

1. Roelfsema, P.R.: Cortical Algorithms for Perceptual Grouping. Annual Review of Neuroscience 29, 203–227 (2006)
2. Hubel, D., Wiesel, T.: Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. The Journal of Physiology 160, 106–154 (1962)
3. Kobatake, E., Tanaka, K.: Neuronal Selectivities to Complex Object Features in the Venral Visual Pathway of the Macaque Cerebral Cortex. Jounal of Neurophysiology 71, 856–867 (1994)
4. Logothetis, N.K., Sheinberg, D.L.: Visual Object Recognition. Annual Review of Neuroscience 19, 577–621 (1996)
5. Tanaka, K.: Inferotemporal Cortex and Object Vision. Annual Review of Neuroscience 19, 109–139 (1996)

6. Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., Fried, I.: Invariant Visual Representation by Single Neurons in the Human Brain. Nature 435(7045), 1102–1107 (2005)
7. Fukushima, K.: Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position. Biological Cybernetics 36(4), 193–202 (1980)
8. Perrett, D.I., Oram, M.: Neurophysiology of Shape Processing. Image and Vision Computing 11, 317–333 (1993)
9. Wallis, G., Rolls, E.T.: A Model of Invariant Object Recognition in the Visual System. Progress in Neurobiology 51, 167–194 (1996)
10. Riesenhuber, M., Poggio, T.: Hierarchical Models of Object Recognition in Cortex. Nature Neuroscience 2(11), 1019–1025 (1999)
11. Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., Poggio, T.: Robust Object Recognition with Cortex-like Mechanisms. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(3), 411–426 (2007)
12. Hyvarinen, A., Hoyer, P.O.: Topographic Independent Component Analysis as a Model of V1, Organization and Receptive fields. Neurocomputing 38, 1307–1315 (2001)
13. Olshausen, B.A., Field, D.J.: Emergence of Simple-cell Receptive Field Properties by Learning a Sparse Code for Natural Images. Nature 381, 607–609 (1996)
14. Inza, I., Larranaga, P., Etxeberria, R., Sierra, B.: Feature Subset Selection by Bayesian Network-based Optimization. Artificial Intelligence 123(1-2), 157–184 (2000)
15. Kohavi, R., John, G.H.: Wrappers for Feature Subset Selection. Artificial Intelligence 97(1-2), 273–324 (1997)
16. Kwak, N., Choi, C.H.: Input Feature Selection by Mutual Information Based on Parzen Window. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(12), 1667–1671 (2002)
17. Oh, I.S., Lee, J.S., Moon, B.R.: Hybrid Genetic Algorithms for Feature Selection. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(11), 1424–1437 (2004)
18. Verikas, A., Bacauskiene, M.: Feature Selection with Neural Networks. Pattern Recognition Letters 23(11), 1323–1335 (2002)

Scheme Implement for Rendering of Chinese Paintings Ink Fluid Style

Tianding Chen

Institute of Communications and Information Technology,
Zhejiang Gongshang University, Hangzhou 310018
chentianding@163.com

Abstract. It proposes a method to view a Chinese landscape painting from an arbitrary viewpoint and synthesizes Chinese ink painting automatically. We analyze the composition of Chinese landscape painting and define 3D basic primitives of mountain and water. First, the user inputs a reference picture. Then according to the feature parameters of 3D basic primitives specified by the user, our system automatically generates a 3D scene model which is the same as the scene of reference. The algorithm extracts strokes from an existing Chinese ink painting image and re-paints it in other styles whatever we want. Using this method we can see resulting images without really performing the painting process. Moreover, the user even requires no painting technique to finish the job. The system contains three stages: modeling, styling and rendering. Furthermore, we integrate the Chinese brush stroke model into our system and have some modifications to achieve Chinese painting styles. So, this research can apply to computer games or virtual environment systems.

1 Introduction

Landscape is essential in Chinese painting. Painters use various kinds of brush styles to depict natural scenery. Terrain is the major object in Chinese landscape painting. Some mountains rise high and erect, and some are gradual and continuous. So the simulation of wrinkles on the rocks becomes very important. Chinese artists capture the spirit and true beauty of the landscape to paint a Chinese landscape painting.

Computer graphics-related research has focused on obtaining photorealistic images. Over many years, the advocate of photorealism has conceived many methods and algorithms for synthetic image generation; the dominant form is based on the physical technology. These are discovered by physical world observation driving, in the light interaction by the surface and the object in the environment. The ray tracing and the radiosity are two powerful technologies in the photorealistic rendering. In both cases, the physical behavior of light is imitated to produce the best photorealism example in a hypothesized world.

Today, it is easy and powerful to construct a photo-realistic virtual world through many developed methods and graphic accelerator. However, photorealism is occasionally not the most effective means of visually expressing emotions. Accordingly, photographs can never entirely replace paintings. Non-photorealistic rendering (NPR) approaches have recently received renewed interest. Researchers have begun to study

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 293 – 299, 2008.
© Springer-Verlag Berlin Heidelberg 2008

how a photograph or a photorealistic image may be made to look like a painting. In recently years, most research in NPR focused on Western painting. Many researches have addressed Western painting, including watercolors, impressionistic painting, pencil sketches and hatching strokes. Lots of painting-algorithms have been proposed for convert a photorealistic image into an art form such as an oil painting or watercolor. Recent research on non-photorealistic rendering has focused on modeling traditional artistic media and styles including pen-and-ink illustration and watercolor painting etc. These methods deliver good results for Western painting. But these approaches are not appropriate to Chinese ink painting. Generally, Western painting involves more precision but Chinese ink painting is more abstract. To simulate the style of Chinese ink painting is not trivial at all. It usually uses brushes and ink as mediums, values the expression of the artistic conception far beyond the precise appearance of the painted subjects. By means of the blending effects of brushes and ink, a painter communicates their frame of mind to the viewers. In contrast to photorealism, in which the driving force is the modeling of physical processes and behavior of light, the processes of human perception can drive NPR techniques. This can be just as demanding as physical simulation but for different reasons: a technologist may find comfort in developing algorithms to reproduce a certain physical phenomenon that is objective and relatively predictable. Developing techniques to replace, augment or even assist the subjective processes of an artist requires a shared understanding of the use to which these techniques will be put. The finest examples of NPR work will be produced when artists and technologists work together to identify a problem and develop solutions that are sympathetic to the creative processes [1].

2 Overview of Rendering Processing

Landscapes have been the main theme of Chinese painting for over one thousand years. It is a form of non-photorealistic rendering. In Chinese landscape painting, rock textures convey the orientation of mountains and contribute to the atmosphere. Several landscape-painting skills are required to capture various types of mountain and water. Over the centuries, masters of Chinese landscape painting have developed various texture strokes. The mature texture strokes may be divided into three groups: line, dot and plane. Chinese artists use many types of brushstrokes to depict nature. A painted of Chinese landscapes must understand both texture strokes and ink brush techniques. A 3D mountain or water is drawn as an outline and texture strokes, using information on the shape, shade and orientation of the object's polygonal surface. This work also uses vertical or slanted brush strokes for drawing outlines and object textures. The main contribution of this work is on the modeling and implementation of an integrated framework for object texture rendering using traditional brush techniques in Chinese landscape painting. The process of producing texture strokes, which includes several basic elements:

(1) 3D information extraction: 3D terrain models rendered using OpenGL include vertices, edges, faces and so on. The luminance map is specified as ink tone values.
(2) Control line construction: a direction field on the surface is computed, silhouette lines are detected and streamlines in 3D object space are generated.

(3) Projection: all control lines, silhouette lines and streamlines are projected onto a 2D viewing plane.
(4) Brush stroke: brush strokes are applied as outlines in the drawing, and the texture strokes are vertical or slanted strokes [2] following streamlines using a rich ink tone specified by a luminance map.
(5) Ink Diffusion: the motion of ink on rice paper is simulated [3].

3 Models and Phenomena for Effects

In order to obtain desired styles, painters need practiced skills during painting. They must have strong domain knowledge about what effects will be produced on the paper with the specific brush stroke and concentrated paint amount. They may not exactly know the physical behaviors or formulas behind the phenomena for effects, but after long periods of practice and countless trial-and-errors, the painters have sufficient experiences to handle these skills to produce desired styles and expressions. Unlike painters, researchers in physics and chemistry are not concerned with the creation of painting styles, but they want to use physical or chemical formulas to describe the underlying behavior behind the specific phenomena. They can therefore make better and cheaper paint to improve paint quality and reduce the cost. Like their fellow researchers in physics and chemistry, when computer scientists or programmers want to make a painting system, the importance things are how to use computational functions to simulate realistic paint motions and behaviors, and then the digital effects can be reproduced. Finally, the end users, digital painters, can transfer their skills in traditional tools and media to virtual tools of the painting system seamlessly.

3.1 Paint Properties

Media in the proposed painting system consist of water particles and several kinds of ink particles (pigments). Unlike others only have one kind of particles for monochrome paintings or three that represent R G B respectively for color painting[2], every pigment in the proposed painting system has a unique identification and its characteristics are like [4]. They behave differently in the simulation by their physical properties and simulation results need to be transformed to R G B in during rendering using their optical properties.

Table 1. Paint Properties

Properties	Water	Pigments
Density	1	2~9
Diameter	$5m\mu$	$50m\mu$~5μ
Absorption		K
Scattering		S
Reflectance		Variable by thickness
Transmittance ratio		Variable by thickness

3.2 Models

Geometric shape of brushes in our painting system can be 2D stamp or 3D polygonal mesh. Shapes can be rigid or deformed by inner brush skeleton dynamics. Although the real brush bristles are volumetric structures that have numerous hair layers from the center axis to surface, brush model in our system coats only two 2D layers in order to convenient computing. The outer layer is called surface layer that contacts with paper or palette surface to perform paint transfer. The inner layer is called reservoir layer, like its literal meaning, which has the ability to preserve more paint than surface layer. It can either receive paint from surface layer or support paint continuously to surface layer.

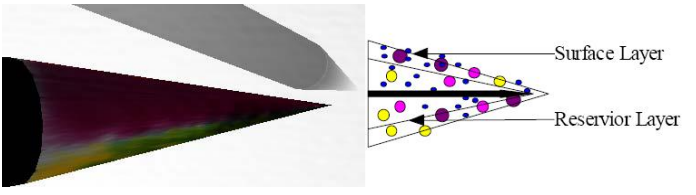


Fig. 1. Brush Polygonal meshes and layer model

In previous works, paper model of ink painting are usually modeled using only one layer. But the knowledge we gain from experiments in last chapter, it is necessary to separate paint that on the paper surface and inside the paper.

The surface layer contains water and ink particles that have not enter the paper yet. We model the inside of the paper using two layers just like our brush model. A 2D upper layer and a 2.5D lower layer are used to approximate the real volumetric paper structure. Upper layer contains active paint flowing between paper cells through connected fiber mesh by capillarity and Brownian motion. Lower layer pigments are already deposited and fixed in paper cells and can not move anymore. The pigments become a permanent part of paper cell structure. Lower layer is a stack of several pigments' reflectance and transparency and thickness. This lower layer is modeled for a main skill call ink accumulation method in Chinese ink painting or impasto in western oil or watercolor painting.

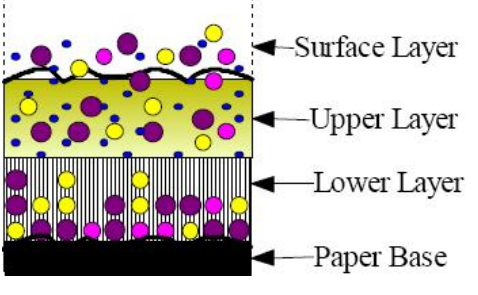


Fig. 2. Paper layer model

3.3 Phenomena

After defining paint properties and field models, this section describes the painting phenomena in our system. Every phenomenon consists of several behaviors described by physical rules and formulas. Paint phenomena are divided into two kinds: "layer-to-layer phenomenon" and "in-layer phenomena." The amount of paint passing from one layer to another is first calculated and recorded in the interface between two layers, and then one of the two adjacent layers subtracts the interface paint and the other adds the paint to satisfy conservation of quantities. Flow-out and flow-in map works like interface; flow-out map record the paint of every paper cell inside one layer flowing to four or eight neighbors and flow-in map record paint flowing from them. New paint amount of every paper cell is computed by adding flow-in paint summation and subtracting flow-out paint summation. See Figure 3 for a illustration for interface and the flow maps.

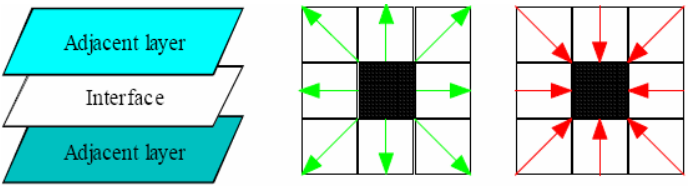


Fig. 3. Interface (Left), flow-out and flow-in map (right)

All phenomenon processes consists water pass and pigment pass; pigments in some phenomenon can only be moved when water move.

4 Painting Styles Algorithm and Result

The first algorithm query the status by drawing a line strips surrounding the tile boundary, if the results are greater then zero, the system will turn on surrounding neighborhood. This algorithm is simple and just need the few memory sizes to save query class and their results. But it turns on all four neighbors at the same time and then neighbor's boundary query results will also turn on their neighbors, some unnecessary updated tiles simulates to reduce performance.

The second algorithms divides line strips surrounding tile boundary into four line segments, only draws and queries line segments if another side tiles are not turned up. This algorithm needs more space to save queries results and may need more drawing calls to query in the beginning, but all the turned up tiles are all necessary to simulate. Then performance will be better than the first one.

Weng's model [5] has 4 parameters to control painting styles: decreasing, concentration, difference and discontinuity. These parameters satisfy the request of simulating calligraphy or plant in Chinese ink painting. However, to present object texture in Chinese landscape painting, we need more parameters to depict effects of dry stroke. In Weng's model, ink quantity and ink color are equivalent, so it is hard to simulate

D_{max}: max distance allowed between two neighboring control points
D_{min}: min distance allowed between two neighboring control points
L_{max}: max length allowed for a stroke
L_{min}: min length allowed for a stroke
L: total length of current stroke s
for every control points c_i in s
    D: distance between c_i and c_{i+1}
    if($D>D_{max}$)
insert new control point(s) between c_i and c_{i+1}
    if($D< D_{min}$)
        remove c_{i+1}
    $L = L + D$
    if($L>L_{max}$)
        terminate this stroke and ***start*** a new stroke
    move to next control point
end of for loop
if($L< L_{min}$)
    reject current stroke

Fig. 4. The algorithm of stroke length control

Fig. 5. With ink painting style simulated result

strokes with dark ink color but low ink quantity. To solve this problem, we use two parameters to present ink quantity and ink color respectively.

When applying strokes, the position, shape and length are decided by control points. All these control points are extracted from object space. So, at each frame, after we project these 3D points to image-space, we can not guarantee the distance between control points. Some control points are close together while others are distant. Besides, the total length of a brush stroke is also uncontrollable. Some strokes look too long, and some are too short. These situations make us difficult to tune up the rendering style of each stroke. Therefore, we propose a mechanism to adjust

distances between control points and strokes' length. If the distance between two control points is far away, we insert additional control points. Otherwise we discard unnecessary control points. If the total length of a stroke is too long, we divide the stroke into several strokes. If the length is too short, we abandon this stroke. The algorithm is shown in Figure 4.

The implementation results of our proposed system are written in C++ language and runs OpenGL on the PC platform with P-Ⅳ 2.4GHz Dual-CPU and 2 GB RAM. It is done by applying a dense brush stroke on the area where you have previously drawn a dilute brush stroke. And the dense stroke must be drawn before the dilute stroke is dried. Figure 5 is simulated result.

5 Conclusions

In this paper, we propose an interactive system for Chinese color ink painting. The proposed system is composed of three main parts: a simple brush model to simulate Chinese writing brush, a method to generate ink diffusion effects in Chinese Ink Painting and a model to represent colors of pigments. The brush model we proposed can produce good Chinese Painting strokes with tablet pen as the input device. By detecting the input pressure of the tablet pen, we can get various strokes with different shapes. The diffusion model is based on real physical phenomenon. Water particles flow in the holes or spaces of fibers of the paper mesh by capillarity. The carbon particles float and move in this liquid due to the collision with water particles. The directions and quantity of water flow are determined as a result of different degree of water absorbency, the alignment of fibers, and the factor of inertia of each paper cell. This ink diffusion algorithm can be used to draw lots of subjects with ink diffusion effect in Chinese Ink Painting style by controlling the user-defined stroke parameters. The most important is the effect of expressing the mixture of different kinds of brush strokes. "Dense brush following dilute brush", is a typical example in Chinese Ink Painting. Besides, since the proposed algorithms are based on physical theory and observational-based analyses, resulting images using this physical-based method are very realistic.

Acknowledgements. This research work is supported by Zhejiang Province Nature Science Foundation under Grant Y107411.

References

1. Green, S.: Introduction to Non-Photorealistic Rendering. SIGGRAPH 1999 Course 17, 2.1-2.7 (1999)
2. Way, D.L., Shih, Z.C.: The Synthesis of Rock Textures in Chinese Landscape Painting. Computer Graphics Forum 20(3), C123–C131 (2001)
3. Huang, S.W., Way, D.L., Shih, Z.C.: Physical-based Model of Ink Diffusion in Chinese Paintings. Journal of WSCG 10(3), 520–527 (2003)
4. Cassidy, J.C., Anderson, S.E., Joshua, E.S., Kurt, W.F., David, H.S.: Computer-generated Watercolor. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, pp. 421–430 (1997)
5. Weng, S.Z., Shih, Z.C., Chiu, H.Y.: The Synthesis of Chinese Ink Painting. In: National Computing Symposium 1999, pp. 461–468 (1999)

Nonnegative Tensor Factorization with Smoothness Constraints

Rafal Zdunek[1] and Tomasz M. Rutkowski[2]

[1] Institute of Telecommunications, Teleinformatics and Acoustics,
Wroclaw University of Technology,
Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland
rafal.zdunek@pwr.wroc.pl
[2] RIKEN Brain Science Institute, Wako-shi, Japan

Abstract. Nonnegative Tensor Factorization (NTF) is an emerging technique in multidimensional signal analysis and it can be used to find parts-based representations of high-dimensional data. In many applications such as multichannel spectrogram processing or multiarray spectra analysis, the unknown features have locally smooth temporal or spatial structure. In this paper, we incorporate to an objective function in NTF additional smoothness constrains that considerably improve the unknown features. In our approach, we propose to use the Markov Random Field (MRF) model that is commonly-used in tomographic image reconstruction to model local smoothness properties of 2D reconstructed images. We extend this model to multidimensional case whereby smoothness can be enforced in all dimensions of a multi-dimensional array. We analyze different clique energy functions that are a subject to MRF. Some numerical results performed on a multidimensional image dataset are presented.

Keywords: Nonnegative Tensor Factorization (NTF), multiarray spectra analysis, Markov Random Field (MRF).

1 Introduction

Nonnegative Tensor Factorization (NMF) [1,2] is an extended version of Nonnegative Matrix Factorization (NMF) [3] to nonnegative multi-dimensional arrays. This method has already found a variety of applications, e.g. in multi-dimensional signal and image processing. Similarly as NMF, NTF also provides lower-rank sparse representations of nonnegative data. Moreover, spatial and temporal correlations between variables can be modeled more accurately with NTF than 2D matrix factorizations. Assuming the variables of interest have all smooth profiles (along each dimension) we can improve NTF with additional smoothness constraints. This approach is justified by observation that usually data representations have a locally smooth spatial and temporal structure, where the locality can be restricted to a few samples.

The smoothness can be modeled in many ways, e.g. by the entropy measure or l_2 norm of the estimated components. In our approach, we used the Markov Random Field (MRF) model that is widely applied in image reconstruction and restoration. Such models, which are often expressed by the Gibbs prior, determine local

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 300–307, 2008.
© Springer-Verlag Berlin Heidelberg 2008

roughness (smoothness) in the analyzed image with consideration of pair-wise interactions among adjacent pixels in a given neighborhood of a singe pixel. Thus, a total smoothness in an image can be expressed by a joint Gibbs distribution with a nonlinear energy function. In our approach, we use the Green's function for measuring strength of the pair-wise pixel interactions. Using a Bayesian framework, we get the Gibbs regularized cost function that is minimized with a gradient descent alternating minimization technique subject to nonnegativity constrains that can be imposed in many ways. One of them is achieved with standard multiplicative updates that were used, e.g. by Lee and Seung [3].

Let $\mathcal{A}$ be a tensor of the order N and dimension $I_1 \times I_2 \times \ldots \times I_{n-1} \times I_n \times I_{n+1} \times \ldots \times I_N$. The n-mode product of the tensor $\mathcal{A}$ and the matrix $\boldsymbol{B} \in \mathbb{R}^{J_n \times I_n}$ is the tensor $\mathcal{A} \times_n \boldsymbol{B}$ of the dimension $I_1 \times I_2 \times \ldots \times I_{n-1} \times J_n \times I_{n+1} \times \ldots \times I_N$ [4]. For $N = 3$, the k-th frontal slice of the tensor $\mathcal{A}$ is the matrix $\boldsymbol{A}_k = \mathcal{A}(:, :, k) \in \mathbb{R}^{I_1 \times I_2}$. Applying the forward cycle matricizing by row-wise unfolding, the tensor $\mathcal{A}$ can be matricized as follows: $\bar{\boldsymbol{A}} = [\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k, \ldots, \boldsymbol{A}_{I_3}] \in \mathbb{R}^{I_1 \times I_2 I_3}$.

The paper has the following organization. The next section introduces the basic factorizations of nonnegative multi-dimensional arrays. Section 3 presents the main contribution, i.e. incorporation of smoothness constraints to the selected models. The simulation experiments are given in Section 5. Finally, some conclusions are drawn in the last section.

2 NTF Models

We assume the observed data are represented by the 3D array $\mathcal{Y} \in \mathbb{R}_+^{I \times T \times K}$, where $\mathbb{R}_+$ denotes the nonnegative orthant (subspace of nonnegative real numbers) of $\mathbb{R}$. There are many models for factorizing the array $\mathcal{Y}$, which follows from the underlying physical phenomena. Typical models [5,6] are as follows:

- **PARAFAC with nonnegativity constraints**

$$\mathcal{Y} = \mathcal{D} \times_1 \boldsymbol{A} \times_2 \boldsymbol{X}^T, \tag{1}$$

  where $\boldsymbol{A} = [a_{ij}] \in \mathbb{R}_+^{I \times J}$, $\boldsymbol{X} = [x_{jt}] \in \mathbb{R}_+^{J \times T}$, and $\mathcal{D} = [d_{jjk}] \in \mathbb{R}_+^{J \times J \times K}$ is a nonnegative tensor with diagonal frontal slices, i.e. $\forall k : \boldsymbol{D}_k = \mathrm{diag}\{d_{jj}\} \in \mathbb{R}^{J \times J}$.
- **NTF 1**

$$\mathcal{Y} = \mathcal{X} \times_1 \boldsymbol{A}, \tag{2}$$

  where $\mathcal{X} = [x_{jtk}] \in \mathbb{R}_+^{J \times T \times K}$ is a nonnegative tensor of lateral sources, and $\boldsymbol{A} = [a_{ij}] \in \mathbb{R}_+^{I \times J}$ is a nonnegative mixing matrix. In practice, NTF 1 is an inexact factorization due to noisy perturbations, i.e.

$$\mathcal{Y} = \mathcal{X} \times_1 \boldsymbol{A} + \mathcal{N}, \tag{3}$$

  where $\mathcal{N} \in \mathbb{R}^{I \times T \times K}$ is a tensor of noisy perturbations.
- **NTF 2**

$$\mathcal{Y} = \mathcal{A} \times_2 \boldsymbol{X}^T + \mathcal{N}, \tag{4}$$

where $\mathcal{A} = [a_{ijk}] \in \mathbb{R}_+^{I \times J \times K}$ is a nonnegative tensor of weighting coefficients, $\boldsymbol{X} = [x_{jt}] \in \mathbb{R}_+^{J \times T}$ is a nonnegative matrix of sources, $\mathcal{N} \in \mathbb{R}^{I \times T \times K}$ is a tensor of noisy perturbations.

The objective is to estimate the multi-dimensional arrays $\mathcal{D}$, $\mathcal{X}$, $\mathcal{A}$, and the matrices $\boldsymbol{A}$ and $\boldsymbol{X}$, subject to nonnegativity constraints of all the entries, given the data tensor $\mathcal{Y}$ and possibly the prior knowledge on the nature of the true components to be estimated or on a statistical distribution of noisy disturbances in $\mathcal{N}$.

In the paper, we assume that the smoothness constraints apply only to the unknown sources, i.e. either to the matrix $\boldsymbol{X}$ or tensor $\mathcal{X}$. The sources may be time- and space-varying with both smooth profiles. The prior knowledge on the sources is particularly strong when the sources have all the profiles locally smooth, i.e. the tensor $\mathcal{X}$ tends to be smooth in all 3 dimensions. This takes place very often when $\mathcal{X}$ models a series of smooth images with a temporal smooth profile. Thus, we restrict our considerations on the smoothness constraints only to the NTF 1 model, since the application of these constraints to the matrix $\boldsymbol{X}$ as in the PARAFAC and NTF 2 is straightforward, and it can be easily achieved. Furthermore, the MRF-based smoothing has been successfully applied to NMF with respect to blind separation of nonnegative signal in [7] and image processing in [8].

3 Smoothness in NTF

To estimate the factors $\boldsymbol{A}$ and $\mathcal{X}$, we use the similar approach as to NMF [3]. The specific cost function $D(\mathcal{Y}||\mathcal{X} \times_1 \boldsymbol{A})$ that measures the distance between $\mathcal{Y}$ and $\mathcal{X} \times_1 \boldsymbol{A}$ is minimized with the alternating minimization technique.

Lee and Seung [3] were the first who proposed two types of NMF algorithms. One minimizes the Euclidean distance, which is optimal for a Gaussian distributed additive noise, and the other for minimization of the Kullback-Leibler (KL) divergence, which is suitable for a Poisson distributed noise. The NMF algorithms that are optimal for many other distribution of additive noise can be found, e.g. in [9,10,11].

Assuming a Poisson noise, which often occurs in image processing, $D(\mathcal{Y}||\mathcal{X} \times_1 \boldsymbol{A})$ is given by the KL divergence. Thus

$$D(\mathcal{Y}||\mathcal{X} \times_1 \boldsymbol{A}) = \sum_{itk} y_{itk} \log \frac{y_{itk}}{z_{itk}} + z_{itk} - y_{itk}, \tag{5}$$

where $\mathcal{Y} = [y_{itk}]$ and $\mathcal{Z} = [z_{itk}] = \mathcal{X} \times_1 \boldsymbol{A}$.

Using the Gibbs prior [12] to model total smoothness in $\mathcal{X}$ and with the Bayesian framework as in [7,8], we get the penalized KL divergence

$$D(\mathcal{Y}||\mathcal{X} \times_1 \boldsymbol{A}) = \sum_{itk} y_{itk} \log \frac{y_{itk}}{z_{itk}} + z_{itk} - y_{itk} + \beta U(\mathcal{X}), \tag{6}$$

where β is a penalty parameter and $U(\mathcal{X})$ is a total energy function that measures total roughness (smoothness) in $\mathcal{X}$.

3.1 Markov Random Field Model

MRF models have been widely applied in many image reconstruction applications, especially in tomographic imaging. In our application, MRF models motivate the definition of the total energy function in (6). Thus

$$U(\mathcal{X}) = \sum_{jtk} \sum_{p \in S^{(j)}} \sum_{q \in S^{(t)}} \sum_{r \in S^{(k)}} w_{pqr}^{(jtk)} \psi\left(x_{jtk} - x_{pqr}, \delta\right), \tag{7}$$

where $S^{(j)}$, $S^{(t)}$, and $S^{(k)}$ are sets of indices of the entries in the neighborhood of x_{jtk}, $w_{pqr}^{(jtk)}$ is a weighting factor between the entries x_{jtk} and x_{pqr}, δ is a scaling factor, and $\psi\left(\xi, \delta\right)$ is some potential function of ξ and δ, which can take various forms. Exemplary potential functions are listed in Table 1

Table 1. Potential functions

Author(s) (Name)	Reference	Functions: $\psi(\xi, \delta)$
(Gaussian)	[12,13]	$\left(\frac{\xi}{\delta}\right)^2$
Besag (Laplacian)	[13]	$\left\lvert\frac{\xi}{\delta}\right\rvert$
Hebert and Leahy	[14]	$\delta \log\left[1 + (\frac{\xi}{\delta})^2\right]$
Geman and McClure	[15]	$\frac{16}{3\sqrt{3}} \frac{(\xi/\delta)^2}{(1 + (\xi/\delta)^2)}$
Geman and Reynolds	[16]	$\frac{\lvert\xi/\delta\rvert}{1 + \lvert\xi/\delta\rvert}$
Stevenson and Delp (Hubert)	[17]	$\min\{\lvert\frac{\xi}{\delta}\rvert^2, 2\lvert\frac{\xi}{\delta}\rvert - 1\}$
Green	[18]	$\delta \log[\cosh(\xi/\delta)]$

The sets $S^{(j)}$, $S^{(t)}$ and $S^{(k)}$, and the associated weighting factors $w_{pqr}^{(jtk)}$ are usually defined by the MRF model. Taking into account the nearest neighborhood, we have $S^{(j)} = \{j-1, j, j+1\}$, $S^{(t)} = \{t-1, t, t+1\}$, and $S^{(k)} = \{k-1, k, k+1\}$. In consequence, $w_{pqr}^{(jtk)} = 1$ for pixels adjacent along a horizontal or vertical line, and $w_{pqr}^{(jtk)} = \frac{1}{\sqrt{2}}$ for pixels adjacent along a diagonal line. For $p = j$, $q = t$, $r = k$, and otherwise, $w_{pqr}^{(jtk)} = 0$.

3.2 Algorithm

Assuming the local stationarity of (6) with respect to $\boldsymbol{A}$ and $\mathcal{X}$, and using the alternating minimization algorithm, NTF 1 can be obtained with the following algorithm:

Set Randomly initialize with nonnegative real numbers: $\boldsymbol{A}^{(0)}$, $\mathcal{X}^{(0)}$,

$$\bar{\boldsymbol{Y}} = [\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_K] \in \mathbb{R}^{J \times TK}, \ \% \textit{ row-wise unfolding}$$

For $s = 1, 2, \ldots,$ until convergence **do**

$$\theta_{jtk} = \sum_{p \in S^{(j)}} \sum_{q \in S^{(t)}} \sum_{r \in S^{(k)}} w_{pqr}^{(jtk)} \frac{\partial}{\partial x_{jtk}} \psi \left(x_{jtk} - x_{pqr}^{(s)}, \delta \right) \Big|_{x_{jtk} \leftarrow x_{jtk}^{(s)}},$$

$$x_{jtk}^{(s+1)} = \frac{x_{jtk}^{(s)}}{\sum_{i=1}^{I} a_{ij}^{(s)} + \beta \theta_{jtk}} \sum_{i=1}^{I} \frac{a_{ij}^{(s)} y_{itk}}{\sum_{c=1}^{J} a_{ic}^{(s)} x_{ctk}^{(s)}},$$

$$\bar{\boldsymbol{X}}^{(s+1)} = [\boldsymbol{X}_1^{(s+1)}, \boldsymbol{X}_2^{(s+1)}, \ldots, \boldsymbol{X}_K^{(s+1)}] \in \mathbb{R}^{J \times TK}, \ \% \textit{ unfolding}$$

$$a_{ij}^{(s+1)} = \frac{a_{ij}^{(s)}}{\sum_{z=1}^{TK} \bar{x}_{jz}^{(s+1)}} \sum_{z=1}^{KT} \frac{\bar{y}_{iz} \bar{x}_{jz}^{(s+1)}}{\sum_{c=1}^{J} a_{ic}^{(s)} x_{cz}^{(s+1)}},$$

$$a_{ij}^{(s+1)} \leftarrow \frac{a_{ij}^{(s+1)}}{\sum_{i=1}^{I} a_{ij}^{(s+1)}}, \ \% \textit{ normalization}$$

End

Since the Green's function [18] in Table 1 satisfies all the properties mentioned in [19], i.e. at a constant $\delta > 0$, it is nonnegative, even, 0 at $\xi = 0$, strictly increasing for $\xi > 0$, unbounded, convex, and has bounded first-derivative, we decided to select this function to our tests. Thus

$$\frac{\partial}{\partial x_{jtk}} \psi \left(x_{jtk} - x_{pqr}, \delta \right) = \tanh \left(\frac{x_{jtk} - x_{pqr}}{\delta} \right).$$

4 Numerical Tests

The proposed algorithm has been extensively tested for various benchmarks of nonnegative smooth and sparse signals and images. The exemplary four original 3D arrays of nonnegative smooth and focussed objects are illustrated in Fig. 1. The arrays are stored in 4D tensor $\mathcal{X} \in \mathbb{R}_+^{32 \times 32 \times 32 \times 4}$.

The mixtures $\mathcal{Y} \in \mathbb{R}_+^{32 \times 32 \times 32 \times 8}$ are obtained by multiplying the tensor $\mathcal{X}$ across 4-mode with the uniformly distributed random matrix $\boldsymbol{A} \in \mathbb{R}_+^{8 \times 4}$ across the 2-nd dimension.

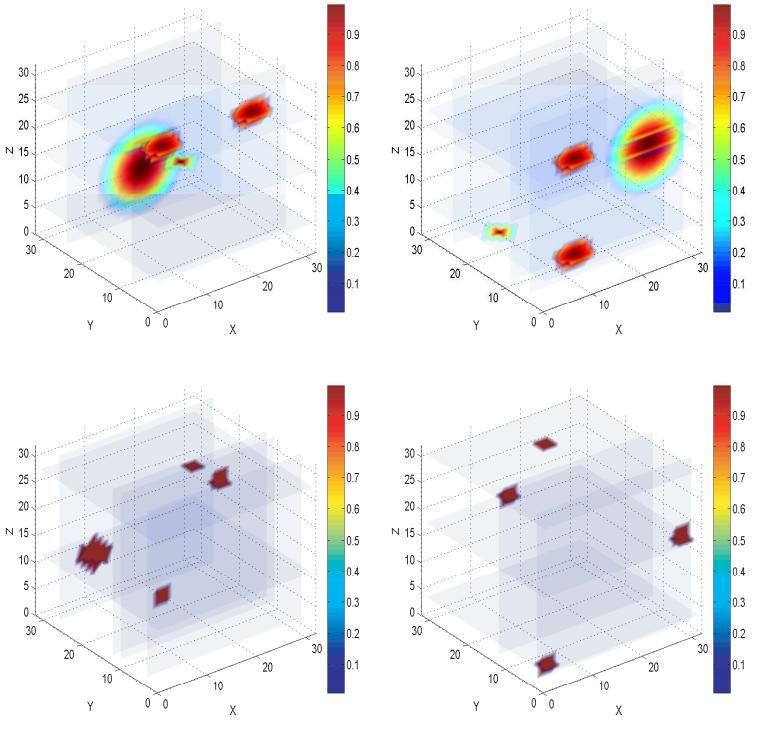


Fig. 1. Volumetric slice plots of 4 original 3D arrays of nonnegative smooth and focussed objects

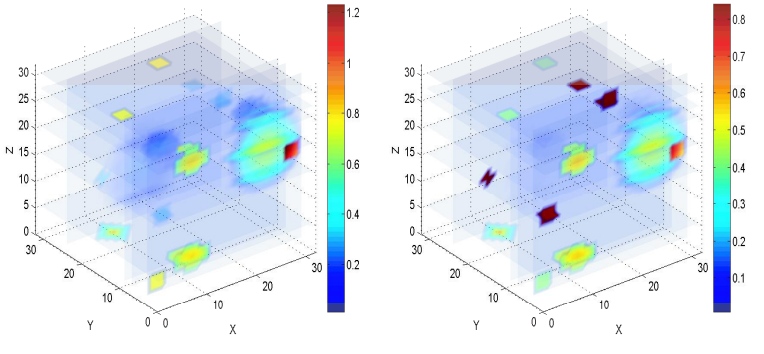


Fig. 2. Volumetric slice plots of 2 selected 3D arrays from 4D tensor of linear mixtures

To separate the original 3D arrays from the mixtures, we used the algorithm presented in Section 3.2. For $\beta = 0$, the algorithm becomes a standard row-unfolded algorithm for NTF 1 [5]. The quality of the separation is evaluated with the standard mean-SIR measure, comparing the estimated 3D arrays with

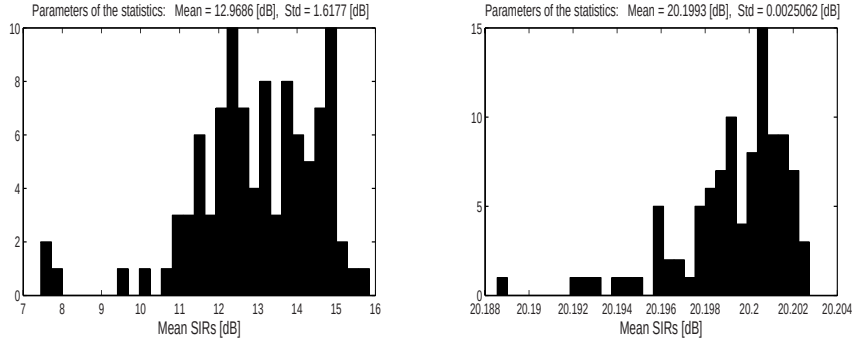


Fig. 3. Histograms of 100 mean-SIR samples generated with: (left) standard algorithm for NTF 1 ($\beta = 0$), (right) our algorithm with the Green's function, $\beta = 0.2$, and $\delta = 0.001$

the respective original ones. The algorithm is initialized with the random initial approximations for $\boldsymbol{A}^{(0)}$ and $\mathcal{X}^{(0)}$. Hence, the Monte Carlo (MC) analysis is applied to test the consistency (repeatability) of the results. Fig. 3 shows the histograms from 100 mean-SIR samples for both $\beta = 0$ and $\beta = 0.2$.

The results demonstrate that the smoothness constraints considerably improve the repeatability of the results (the STD of the histogram with $\beta = 0.2$ is much smaller than with $\beta = 0$). This is a promising result towards a uniqueness of NTF.

5 Conclusions

In the paper, we derived the new algorithm for NTF 1, which may be useful for estimation of locally smooth signals and images in BSS applications. The algorithm exploits the information on pair-wise interactions between adjacent pixels, which is motivated by MRF models in tomographic image reconstruction. The proposed approach can be further extended with additional constraints or different updating rules. Also, another extension may concern the application of data-driven hyperparameter estimation techniques, especially for the regularization parameter.

References

1. Shashua, A., Hazan, T.: Non-negative tensor factorization with applications to statistics and computer vision. In: Proc. of the 22th International Conference on Machine Learning, Bonn, Germany (2005)
2. Hazan, T., Polak, S., Shashua, A.: Sparse image coding using a 3D non-negative tensor factorization. In: International Conference of Computer Vision (ICCV), pp. 50–57 (2005)
3. Lee, D.D., Seung, H.S.: Learning the parts of objects by nonnegative matrix factorization. Nature 401, 788–791 (1999)

4. Bader, B.W., Kolda, T.G.: Algorithm 862: Matlab tensor classes for fast algorithm prototyping. ACM Trans. Math. Softw. 32, 635–653 (2006)
5. Cichocki, A., Zdunek, R., Choi, S., Plemmons, R., Amari, S.I.: Novel multi-layer nonnegative tensor factorization with sparsity constraints. In: Beliczynski, B., Dzielinski, A., Iwanowski, M., Ribeiro, B. (eds.) ICANNGA 2007. LNCS, vol. 4432, pp. 271–280. Springer, Heidelberg (2007)
6. Cichocki, A., Zdunek, R., Choi, S., Plemmons, R., Amari, S.: Nonnegative tensor factorization using Alpha and Beta divergencies. In: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2007), Honolulu, Hawaii, USA, vol. III, pp. 1393–1396 (2007)
7. Zdunek, R., Cichocki, A.: Gibbs regularized nonnegative matrix factorization for blind separation of locally smooth signals. In: 15th IEEE International Workshop on Nonlinear Dynamics of Electronic Systems (NDES 2007), Tokushima, Japan, pp. 317–320 (2007)
8. Zdunek, R., Cichocki, A.: Blind image separation using nonnegative matrix factorization with Gibbs smoothing. In: Ishikawa, M., Doya, K., Miyamoto, H., Yamakawa, T. (eds.) ICONIP 2007, Part II. LNCS, vol. 4985, pp. 519–528. Springer, Heidelberg (2008)
9. Dhillon, I., Sra, S.: Generalized nonnegative matrix approximations with Bregman divergences. In: Neural Information Proc. Systems, Vancouver, Canada, pp. 283–290 (2005)
10. Cichocki, A., Zdunek, R., Amari, S.: Csiszar's divergences for non-negative matrix factorization: Family of new algorithms. In: Rosca, J.P., Erdogmus, D., Príncipe, J.C., Haykin, S. (eds.) ICA 2006. LNCS, vol. 3889, pp. 32–39. Springer, Heidelberg (2006)
11. Kompass, R.: A generalized divergence measure for nonnegative matrix factorization. Neural Computation 19, 780–791 (2006)
12. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6, 721–741 (1984)
13. Besag, J.: Toward Bayesian image analysis. J. Appl. Stat. 16, 395–407 (1989)
14. Hebert, T., Leahy, R.: A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors. IEEE Transactions on Medical Imaging 8, 194–202 (1989)
15. Geman, S., McClure, D.: Statistical methods for tomographic image reconstruction. Bull. Int. Stat. Inst. LII-4, 5–21 (1987)
16. Geman, S., Reynolds, G.: Constrained parameters and the recovery of discontinuities. IEEE Trans. Pattern Anal. Machine Intell. 14, 367–383 (1992)
17. Stevenson, R., Delp, E.: Fitting curves with discontinuities. In: Proc. 1st Int. Workshop on Robust Computer Vision, Seattle, Washington, USA (1990)
18. Green, P.J.: Bayesian reconstruction from emission tomography data using a modified EM algorithm. IEEE Trans. Medical Imaging 9, 84–93 (1990)
19. Lange, K., Carson, R.: EM reconstruction algorithms for emission and transmission tomography. J. Comp. Assisted Tomo. 8, 306–316 (1984)

A Minimum Zone Method for Evaluating Straightness Errors Using PSO Algorithm

Ke Zhang

School of Mechanical and Automation Engineering
Shanghai Institute of Technology, 200235 Shanghai, China
zkwy@hotmail.com

Abstract. Straightness error is the commonest one of form errors. In this paper, an intelligent evaluation method for straightness errors is provided. According to characteristics of straightness error evaluation, Particle Swarm Optimization (PSO) is proposed to evaluate the minimum zone error. The evolutional optimum model and the calculation process for using the PSO to evaluate minimum zone error are introduced in detail. Compared with conventional optimum methods such as simplex search and Powell method, PSO algorithm can find the global optimal solution, and the precision of calculating result is very good. Compared to genetic algorithms (GA), PSO is easy to implement and there are few parameters to adjust. Finally, a numerical example is carried out. The control experiment results evaluated by using different method such as the least square, simplex search, Powell optimum methods, Simulated Annealing Algorithms (SAA), GA and PSO, indicate that the proposed method does provide better accuracy on straightness error evaluation, and it has fast convergent speed as well as using computer.

Keywords: Genetic Algorithms (GA),Evaluating Straightness, PSO Algorithm.

1 Introduction

The minimum zone deviation, as defined in the ISO/R1101 standard, is generally accepted for specifying the form errors of geometric features [1,2]. However, no specific methods are recommended for finding minimum zone deviation in ISO/R1101. At present, the evaluation methods of form and position error are the least square, minimum zone method and so on. Although the least-squares method, because of its simplicity in computation and uniqueness of the solution provided, is most widely used in industry for determining form and position error, it provides only an approximate solution that does not guarantee the minimum zone value [3]. The results of minimum zone method not only verge on ideal error value, but also accord with ISO standard. Therefore, much research has been devoted to finding the minimum zone solutions for straightness error and other form errors using a variety of methods. Some researchers applied the numerical methods of linear programming [3,4,5], such as the Monte Carlo method, the simplex search, spiral search, and the minimax approximation algorithm etc. Another approach has been to find the enclosing polygon for the minimum zone solution, such as the eigen-polyhedral method, the convex polygon method, and the convex hall theory etc.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 308 –314, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The optimization algorithms are commonly used to approach the minima of straightness error objective function through iteration when a microcomputer is applied to assess straightness errors by minimum zone method. The essential prerequisite for convergence of any optimization algorithm is that the objective function to be solved has only one minimum in its definition domain, that is, it is a single valley one. If an objective function has more local minima in its definition domain, it solution searched for by an optimization algorithms may not be its global minimum which is the wanted straightness error. Therefore, the mathematical models and algorithms for straightness error evaluation may be influenced in their solutions' reliability and practical value.

The traditional optimization methods are employed to refine the least-squares solution further. However, these traditional optimization methods have drawbacks in finding the global optimal solution, because it is so easy for these traditional methods to trap in local minimum points [6].

Particle Swarm Optimization (PSO) is an evolutionary computation technique developed by Dr. Eberhart and Dr. Kennedy in 1995 [7, 8, 9], inspired by social behavior of bird flocking or fish schooling. It is designed for nonlinear functions. It can find the solution to the problems of optimization with multi-variables and multi-parameters. PSO is inspired by the behavior of bird flocking. Particle swarm optimization (PSO) as one of the modern heuristic algorithms is also a population based evolutionary algorithm. Unlike population based evolutionary algorithms, PSO is motivated by the simulation of social behavior instead of survival of the fittest, and each candidate solution is associated with a velocity. The candidate solutions, called "Particles" then "fly" through the search space. Similar to genetic algorithms (GA), PSO is a population based optimization tool. The system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike GA, PSO has no evolution operators such as crossover and mutation. Compared to GA, the advantages of PSO are that PSO is easy to implement and there are few parameters to adjust. PSO has been successfully applied in many areas: function optimization, artificial neural network training, fuzzy system control, and other areas [10, 11, 12] where GA can be applied.

In this paper, based on characteristics of straightness error evaluation, a novel algorithm PSO is presented to find the value of straightness error according to the minimum zone criterion. The rest of this paper is organized as follows: The Particle Swarm Optimization is given in Section 2. The straightness error analysis is presented in Section 3. An example is given to test its validity in Section 4. Simulation results and the discussion of the results are presented therein. Finally, conclusions are given in Section 5.

2 Particle Swarm Optimization

Swarm intelligence is an emerging field of biologically-inspired artificial intelligence based on the behavioral models of social insects such as ants, bees, wasps and termites. This approach utilizes simple and flexible agents that form a collective intelligence as a group. Since 1990s, swarm intelligence has already become the new research focus and swarm-like algorithms, such as PSO, have already been applied

successfully to solve real-world optimization problems in engineering and telecommunication. The PSO algorithm does not use the filtering operation (such as crossover and mutation) and the members of the entire population are maintained through the search procedure.

2.1 Particle Swarm Optimization

PSO simulate social behavior, in which a population of individuals exists. These individuals (also called "particles") are "evolved" by cooperation and competition among the individuals themselves through generations [12,13,14]. In PSO, each potential solution is assigned a randomized velocity, are "flown" through the problem space. Each particle adjusts its flying according to its own flying experience and its companions' flying experience. The *i*th particle is represented as $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$. Each particle is treated as a point in a D-dimensional space. The best previous position (the best fitness value is called *pBest*) of any particle is recorded and represented as $P_i = (p_{i1}, p_{i2}, \ldots, p_{id})$. Anther "best" value (called *gBest*) is recorded by all the particles in the population.

This location is represented as $P_g = (p_{g1}, p_{g2}, \ldots, p_{gD})$. At each time step, the rate of the position changing velocity (accelerating) for the *i*th particle is represented as $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$. Each particle moves toward its *pBest* and *gBest* locations. The performance of each particle is measured according to a fitness function, which is related to the problem to be solved. The particles are manipulated according to the following equation:

$$\begin{aligned} v_{id} &= w \cdot v_{id} + c_1 \cdot rand(\ ) \cdot (p_{id} - x_{id}) \\ &\quad + c_2 \cdot Rand(\ ) \cdot (p_{gd} - x_{id}) \end{aligned} \tag{1}$$

$$x_{id} = x_{id} + v_{id} \tag{2}$$

where c_1 and c_2 in equation (1) are two positive constants, which represent the weighting of the stochastic acceleration terms that pull each particle toward *pBest* and *gBest* positions. Low values allow particles to roam far from the target regions before being tugged back. On the other hand, high values cause abrupt movement toward, or past, target regions. Proper values of c_1 and c_2 can quicken convergence. *rand( )* and *Rand( )* are two random functions in the range [0, 1]. The use of the inertia weight *w* provides a balance between global and local exploration, and results in less iteration to find an optimal solution.

2.2 Parameters Setting

Main parameters of PSO are set as follows.

(1) The inertia parameter

The inertial parameter *w* is introduced to control the impact of the previous history of velocities on current velocity. A larger inertia weight facilitates global optimization while smaller inertia weight facilitates local optimization. In this paper, the inertia

weight ranges in a decreasing way in an adaptive way. The inertia weight is obtained by the following equation:

$$W = W_{max} - \frac{W_{max} - W_{min}}{Iter} \cdot iter \tag{3}$$

Where W_{max} : the maximum value of W

W_{min} : the minimum value of W

Iter: the maximum iteration number of PSO

iter: current iteration number of PSO

(2) The parameters c_1 and c_2

The acceleration constants c_1 and c_2 control the maximum step size the particle can do, they are set to 2 according to experiences of other researchers.

2.3 Optimization Algorithm

The procedure of PSO algorithm is shown as follows.

Step1. Initialize ***V*** and ***X*** of each particle.

Step2. Calculate the fitness of each particle .

Step3. If the fitness value is better than the best fitness value (*pBest*) in history, set current value as the new *pBest*.

Step4. Choose the particle with the best fitness value of all the particles as the *gBest* .

Step5. Calculate particle velocity according equation (1), update particle position according equation (2).

Step6. When the number of maximum iterations or minimum error criteria is attained, the particle with the best fitness value in *X* is the approximate solution and STOP; otherwise let *iter*=*iter*+1 and turn to Step 2. Where *iter* denotes the current iteration number.

Particles' velocities on each dimension are clamped to a maximum velocity $v_{\max}$, If the sum of accelerations would cause the velocity on that dimension to exceed $v_{\max}$, which is a parameter specified by the user. Then the velocity on that dimension is limited to $v_{\max}$.

3 Straightness Analysis

For an element of a surface, straightness error is measured by finding two parallel lines minimum distance apart that contain the element（see in Fig.1）. Mathematically, this may be defined as follows. Given a set *S* of measured points, find two parallel lines minimum distance apart that contain the set *S*. The two parallel lines define the smallest feasible region within which set *S* must fall.

Such two-parallel lines can be represented by

$$\begin{cases} y = Ax + B_1 \\ y = Ax + B_2 \end{cases} \tag{4}$$

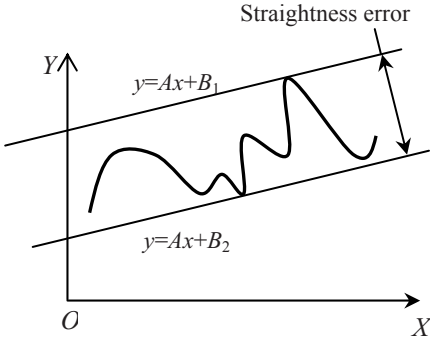


Fig. 1. Straightness by minimum zone method

where A, B_1, and B_2 are coefficients. If x-, y- coordinates are obtained, both B_1 and B_2 become a function of A. Define the measured data (x_i, y_j), where $i = 1,2,\ldots,n$; n is number of data. Therefore, the distance t between such planes is

$$\begin{aligned} t &= \frac{|B_1 - B_2|}{(1+A)^{1/2}} \\ &= \frac{\max_i(y_i - Ax_i) - \min_i(y_i - Ax_i)}{(1+A^2)^{1/2}} \end{aligned} \tag{5}$$

which is a function of A. The objective function is defined as follows, by considering the distance between the two parallel straightness lines that contain the straightness profile, as shown in Fig.1.

$$f = \frac{\max_i(y_i - Ax_i) - \min_i(y_i - Ax_i)}{(1+A^2)^{1/2}} \tag{6}$$

where (x_i, y_i) are measured data, A is optimization variable. Therefore, the minimum zone straightness is expressed as follows:

$$\min f(A) \tag{7}$$

The exact value of straightness error may be found by solving the above optimization problem.

4 Examples

One numerical example is given here to check the validity and efficiency of the proposed evaluation method. A set of data points measured is shown in Table 1 [5]. The procedures were programmed in the Matlab programming language. The calculation result of straightness using the hybrid algorithm is $f = 5.025\,\mu m$. The results of straightness evaluation from different methods are provided in Table 2. As shown in

Table 1. Data Measured

No.	*x* (mm)	*y* (mm)
1	0.3952	-0.0032
2	0.6953	-0.0016
3	0.9669	-0.0042
4	1.2762	-0.0028
5	1.5797	-0.0037
6	1.8593	-0.0007
7	2.1333	-0.0010
8	2.4197	0.0007
9	2.6001	0.0007
10	2.5890	0.0017
11	3.0662	0.0025
12	3.2165	-0.0017
13	3.4217	0.0026
14	3.6179	0.0027
15	3.8185	0.0047

Table 2. Straightness Results

Calculation method	Straightness error
Least-square	5.377
Simplex search	5.186
Powell search	5.188
Simulated Annealing Algorithm	5.062
Genetic Algorithm	5.038
PSO Algorithm	5.025

Data unit, micron

the table, the comparison that the global optimum solution of straightness evaluation problem using the proposed procedure can be obtained and accord with the fact of measured profile. The results also show the PSO evaluation value is slightly superior to GA and SAA.

5 Conclusions

In this paper, an intelligent optimization approach to evaluate straightness error was presented. The straightness evaluation problem was formulated as unconstraint optimization problems. Solution procedure of Particle Swarm Optimization was developed to solve the optimization problem. The evaluation method was compared to some of existing techniques. It is shown through example that the procedure of this paper provides exact values of straightness error. The result also shows that the proposed procedure converges to the global optimum more rapidly and well than conventional methods. The evaluation method is the same with the others form and position error evaluation.

Acknowledgments. The work is supported in part by the Science Foundation of Shanghai Institute of Technology (No. YJ200609). The authors thank Dr. Wang Hairong and Dr. Lu Dejiang of Xi'an Jiaotong University for giving us some advice.

References

1. Wang, M., Cheraghi, H., Masud, A.S.: Circularity Error Evaluation Theory and Algorithm. Precision Engineering 23(2), 164–176 (1999)
2. Cheraghi, S.H.: Evaluating the Geometric Characteristics of Cylindrical Features. Precision Engineering 27(2), 195–204 (2003)
3. Kanad, T., Suzuki, S.: Evaluation of Minimum Zone Flatness by Means of Nonlinear Optimization Techniques and Its Verification. Precision Engineering 15(2), 93–99 (1993)
4. Huang, S.T., Fan, K.C., Wu, J.H.: A New Minimum Zone Method for Evaluating Flatness error. Precision Engineering 15(1), 25–32 (1993)
5. Cheraghi, S.H., Lim, H.S., Motavalli, S.: Straightness and Flatness Tolerance Evaluation: an Optimization Approach. Precision Engineering 18(1), 30–37 (1996)
6. Singiresu, S.R.: Engineering Optimization. John Wiley & Sons, New York (1996)
7. Eberhart, R., Kennedy, J.: A New Optimizer Using Particle Swarm Theory. In: Proc. of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, pp. 39–43 (1995)
8. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: IEEE Int1. Conf. on Neural Networks, pp. 1942–1948. Perth (1995)
9. Kennedy, J., Eberhart, R.C.: A Discrete Binary Version of the Particle Swarm Algorithm. In: Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics, pp. 4104–4109. IEEE Service Center, Piscataway (1997)
10. Eberhart, R., Shi, Y.: Tracking and Optimizing Dynamic Systems with Particle Swarms. In: Proceedings of the IEEE Congress on Evolutionary Computation, Seoul, pp. 94–97 (2001)
11. Trelea, I.C.: The Particle Swarm Optimization Algorithm: Convergence Analysis and Parameter Selection. Information Processing Letters 6, 317–325 (2003)
12. Kennedy, J., Ceberhart, R., Shi, S.Y.: Swarm Intelligence. Morgan Kaufmann, San Mateo (2001)
13. Clerc, M., Kennedy, J.: The Particle Swarm Explosion Stability and Convergence in a Multidimensional Complex Space. IEEE Transaction on Evolutionary Computation 6(1), 58–73 (2002)
14. Yoshida, H., Kawata, K., Fukuyama, Y., Nakanishi, Y.: A Particle Swarm Optimization for Reactive Power and Voltage Control Considering Voltage Stability. In: Proceedings Int1. Conf. on Intelligent System Application to Power Systems, pp. 117–121 (1999)

A Supervised PLS Information Recognition Algorithm

Fengxiang Jin[1] and Shifei Ding[2]

[1] College of Geoinformation Science and Engineering, Shandong University of Science and Technology, Qingdao 266510 P.R. China
[2] School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221008 China
dingsf@cumt.edu.cn, dingshifei@sina.com

Abstract. This paper analyzes the basic principle of Partial Least Squares Regression (PLS), and based on PLS, a Supervised Information Recognition (SIR) algorithm is set up. The algorithm is built up by regarding response variables as 0-1 variables, and it organically combines feature extraction and classifier design in pattern recognition. Compared with Fisher discriminative analysis, Bayes discriminative analysis and some other classical pattern recognition algorithms, SIR algorithm has a much more powerful information recognition ability, and it doesn't demand the data to have good distribution, especially it is more effective to collinearity information or explaining variables are more but the sample size is small. Applied this algorithm to classification recognition of land quality, the results show that the algorithm established in this paper is effective and reliable.

Keywords: We would like to encourage you to list your keywords in this section.

1 Introduction

Pattern recognition is an important field of intelligent information processing, it is not a simple systematics, it contains the description of the system, comprehension and synthesis, and it needs to learn, decide, and find the rules from complex procedure through large amounts of information. It was firstly advanced in 1920s, along with the emergence of the first computer in 1940s, the rise of artificial intelligence in 1950s, pattern recognition rapidly became a subject in 1960s, the theories and methods that it studies are widely taken into account in many fields of science and technology, it accelerates the development of artificial intelligence[1,2]. In most pattern recognition systems, feature extraction and classifier design are the two most key steps of pattern recognition systems, this directly has respect to the recognition results. In feature extraction, correlation analysis, PCA, factor analysis, ICA and other multivariate statistical methods are widely applied[3,4]. In classifier design, the most widely used methods are Nearest Neighbor Classifier (NNC) and cosine Classifier(cosC)[5]. Obviously, composing different feature extraction methods and different classifiers, the recognition performance will be different. But until now, generally we regard feature extraction and classifier design as two completely independent steps. It is a significant study topic how to organically combine feature extraction and classifier design, so as

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 315 – 320, 2008.
© Springer-Verlag Berlin Heidelberg 2008

to reach the best recognition performance. In the study of machine learning, pattern recognition and classification, we usually encounter the interdependence relationship between two groups of multiplex correlative variables, and in order to study using one group of independent variables to predict the other group of induced variables, a frequently used statistical modeling method is classical Multiple Linear Regression(MLR) method which is based on ordinary least square(OLS). But only when the independent variables meet: ①the number of the variables is few; ②no multicollinearity; ③it is easy to explain the relationship between each independent variable and induced variable, the OLS estimation of MLR has a fine theory characteristic, and can fitting original data better. If the data don't meet the above-mentioned three conditions, OLS estimation will fail, overfitting, regression coefficient or prediction results are hard to explain and other situations will arise[6,7].

Partial least squares regression method emerged and developed in recent years is a multivariate data processing method that based on PCA , PCR, it is widely applicable. It is the further expansion of OLS, was firstly advanced in 1970s by Herman Wold, who is an econometrician, in 1980s, it was applied in the fields such as stoichiometric chemistry, industrial design and so on, and worked well, so the value of PLS method was fully accepted by experts, and it was expanded to other fields[8,9]. PLS is usually applied to soft modeling of data, in the process of modeling, it gathers the characteristic of principal component analysis, classical correlation analysis and linear regression analysis, can provide a more reasonable regression model, besides, it can also achieve some research that is similar to PCA, CCA, provide some more plentiful and deep information. This paper firstly analyzes PCR, points out its advantages and disadvantages, further studies the PLS algorithm, based on this, we advance land quality classification algorithm based on PLS, and compare it with the results of PCA, fully explain the vantage of PLS used in land quality classification.

2 PLS Information Recognition Algorithm

2.1 PLS Modeling Thought

Supposed there're m explaining variables $\{x_1, x_2, \cdots, x_m\}$, p response variables $\{y_1, y_2, \cdots, y_p\}$, after observing n sample points, we construct observing matrix $X = (x_{ij})_{n\times m}$, $Y = (y_{ij})_{n\times p}$ of explaining variables and response variables.

The basic thought of PLS is searching some linear combination in explaining variable space, so as to better explain the variation information of response variables, that is to extract components t_1, u_1 from data matrix X, Y, t_1 is a linear combination of $x_1, x_2, \cdots, x_m$, u_1 is a linear combination of $y_1, y_2, \cdots, y_p$, and it should meet the two following conditions:

① t_1 and u_1 should carry variation information of each data set as much as possible;

②the degree of correlation between t_1 and u_1 should be maximum. If the conditions are satisfied, t_1 and u_1 will include the information of data matrix X, Y to the greatest extent possible, at the same time independent variable component t_1 has the

strongest explanatory ability to the induced variable component u_1. After extracting the first pair of component, we should implement the regression X to t_1 and the regression Y to t_1 separately, if the regression equation meets the default accuracy, then the algorithm stops; otherwise extracting the second pair of component using the residual information of X explained by t_1 and Y explained by t_1, repeating the procedure, until satisfying the demand. If extracting r components $t_1, t_2, \cdots, t_r$ to X, at last implementing the regression of $y_k (k = 1,2,\cdots, p)$ on $t_1, t_2, \cdots, t_r$, then transforming to the regression on $x_1, x_2, \cdots, x_m$, so we finish the PLS regression modeling procedure.

2.2 PLS Information Recognition Algorithm

If response variables (Y) are divided into c levels, then we can transform the observing matrix of response variables to extreme matrix Z, according to the modeling thought of PLS, we can build the PLS model of X and Z, the algorithm procedure is as follows:

Step 1. Extracting the first pair of components from X, Z separately, and making the relativity maximum;

Suppose if the first pair of components that extracted from X, Z are t_1、 u_1, and their score vectors are

$$t_1 = Xw_1 = (t_{i1})_{n\times 1}, \quad u_1 = Zv_1 = (u_{i1})_{n\times 1} . \tag{1}$$

The degree of correlation between t_1 and u_1 can be expressed as their inner product, that is we can get the following optimization problem:

$$\begin{cases} \max S = t_1' u_1 = (Xw_1)'(Zv_1) = w_1' X' Z v_1 \\ s.t. \begin{cases} w_1' w_1 = 1 \\ v_1' v_1 = 1 \end{cases} \end{cases} . \tag{2}$$

Using Lagrange's method of multipliers, we can get the problem's optimum relation, that is, we only need to compute the eigenvalue of matrix $M = X'YY'X$ and the corresponding eigenvector, then M's maximum eigenvalue is the standardized eigenvector corresponding to λ_1^2, and this eigenvalue is w_1, and v_1 can be computed from w_1:

$$v_1 = \frac{1}{\lambda_1} Y'Xw_1 . \tag{3}$$

In PLS, $w_1 = (w_{11}, w_{12}, \cdots, w_{1m})'$ is called model effect weight; and $v_1 = (v_{11}, v_{12}, \cdots, v_{1m})'$ is called induced variable weight.

Step 2. Establishing the regression of Z on t_1 and the regression of X on t_1

Suppose the regression model is

$$\begin{cases} X = t_1\alpha_1' + E_1 \\ Z = t_1\beta_1' + F_1 \end{cases} . \tag{4}$$

In it, $\alpha_1 = (\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1m})'$ and $\beta_1 = (\beta_{11}, \beta_{12}, \cdots, \beta_{1p})'$ are regression coefficient vectors, and E_1 and F_1 are residual vectors, so the least-squares estimation of α_1 and β_1 are

$$\alpha_1 = \frac{X't_1}{\|t_1\|^2}, \beta_1 = \frac{Z't_1}{\|t_1\|^2} . \tag{5}$$

In it, α_1 is model effect loading。

Step 3. Using residual matrix E_1 and F_1 to replace X and Z, repeating the steps above

If $\hat{X} = t_1\alpha_1'$, $\hat{Z} = t_1\beta_1'$, then the residual matrix will be $E_1 = X - \hat{X}$, $F_1 = Z - \hat{Z}$. If the absolute value of the element in residual error F_1 is approximating to zero, then we consider the regression precision that is established by the first component has met the demand, can stop extracting component. Or else we should use residual matrix E_1 and F_1 to replace X and Z, and repeat the procedure above.

Step 4. Suppose if the rank of data matrix X satisfies $R(X) = r \le \min(n-1, m)$, then there are r components $t_1, \cdots, t_r$, caused

$$\begin{cases} X = t_1\alpha_1' + t_2\alpha_2' + \cdots + t_r\alpha_r' + E_r \\ Y = t_1\beta_1' + t_2\beta_2' + \cdots + t_r\beta_r' + F_r \end{cases} . \tag{6}$$

Suppose if $X_i^*(i = 1, \cdots .m), Z^*$ expresses standardized variable, using $t_k = w_{k1}X_1^* + \cdots + w_{km}X_m^*$ to substitute in the second formula of formula(6), then we get the PLS regression mathematical equation of p standardized induced variables

$$\hat{Y}_j^* = a_{j1}X_1^* + a_{j2}X_2^* + \cdots + a_{jm}X_m^* \ (j = 1,2, \cdots, p) . \tag{7}$$

Step 5. Use cross-validation method to determine the number of extracting component l;

Generally speaking, PLS regression doesn't need to use all of the existing r components $t_1, t_2, \cdots, t_r$ to establish regression equation, but like PCA, only use the front l components $(l \le r)$, so we can get the regression equation with better prediction ability. Generally we use cross-validation[6,7], that is to leave k samples, use other training samples to model, in order to predict the left k samples, to compute the prediction residual error, all of the training samples are left in turn, so as to model, predict, compute prediction residual error sum of squares(PRESS), then we get the definition of PRESS, that is

$$PRESS = \sum_{i=1}^{n} (z_{i1} - z_{i1}^{p})^2 . \tag{8}$$

In the formula, z_{i1}^{p} is the prediction value, z_{i1} is the original target value. When k=1, it is so-called leave-one-out method. In practice, we can decide according to the change of PRESS, and with the change trend of fitting residual sum of squares, appoint the number of extracting latent variables with human attempt.

Step 6. Using PLS model to do pattern recognition and classification.

3 Example on Analysis

We chose 17 correlative factors of land quality evaluation related to the productivity of land, they're soil organic matter content(%), total nitrogen content(%), rapidly available phosphorus(ppm), rapidly available potassium(ppm), cation exchange capacity(m1/100g), PH value, texture, water table(m), per mu yield(catty), the original data are omitted. Applied DPS data processing system, the relationship between prediction residual error sum of squares and latent variables is shown in Figure 1.

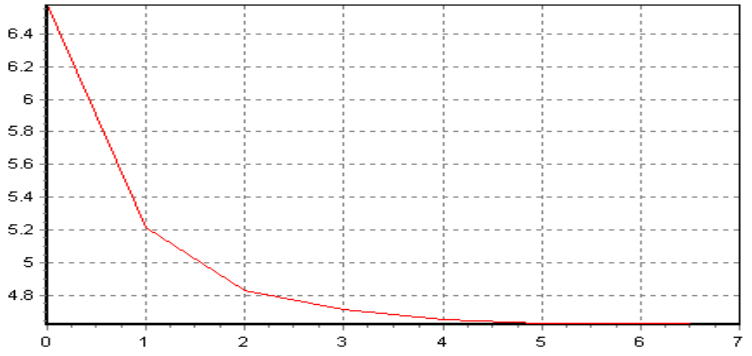


Fig. 1. The relationship between PRESS and latent variable k

According to Fig.1, when PRESS=5.03905, PRESS value is small, and the Change tendency is slow, so we choose the latent variable k=4, and the PLS regression equation is as follows:

$$y = 0.046019 + 0.042429x_1 + 13.228457x_2 + 0.024681x_3 + 0.001044x_4 \\ -0.037289x_5 - 0.075573x_6 + 0.251966x_7 - 0.002759x_8 , \quad (9)$$

the decision accuracy rate of PLS information recognition is 93.33%, the prediction accuracy rate is 100%; the decision accuracy rate of Fisher discriminative is 86.67%, the prediction accuracy rate is 100%; the decision accuracy rate of Bayes discriminative is 93.33%, the prediction accuracy rate is 100%. So the land quality classification based on PLS has a higher accuracy rate, compared with Bayes discriminative analysis, its demand to the sample distribution is lesser, so it is easy to be popularized and applied in practice.

4 Results and Discussion

PLS regression can implement the comprehensive application of several data analysis methods, so it is called the second generation regression method. It gets the basic function of MLR, PCA, CCA together, and organically combines the prediction analysis of modeling type and data exploration analysis of non modeling, so as to fit the goals of different studies.

PLS regression is a robust statistical analysis technique of data soft modeling. When the explaining variables are more but the observation samples are less, it's PLS's vantage. When the observation samples are less than the explaining variables, OLS can't be used as it usually requires that the observation samples should not less than 8-10 times of the explaining variables; but PLS is not restricted by this, it only needs to choose suitable "direction" of explaining variables space, then its fitting and prediction to the data will be robust and reliable.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant No.40574001.

References

1. Duda, R.O., Hart, P.E. (eds.): Pattern Classification and Scene Analysis. Wiley, New York (1973)
2. Bian, Z.Q., Zhang, X.G. (eds.): Pattern Recognition. Tsinghua University Press, Beijng (2000)
3. Ding, S.F., Shi, Z.Z.: Supervised Feature Extraction Algorithm Based on Improved Polynomial Entropy. Journal of Information Science 32(4), 309–315 (2006)
4. Ding, S.F., Shi, Z.Z., Liang, Y.: Information Feature Analysis and Improved Algorithm of PCA[A]. In: Proceedings of 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, pp. 1756–1761 (2005)
5. Ding, S.F., Shi, Z.Z.: Studies on Incidence Pattern Recognition Based on Information Entropy. Journal of Information Science, 2005 31(6), 497–502 (2005)
6. Rao, C.R., Toutenburg, H.: Linear model. Springer, New York (1995)
7. Wang, H.W.: Partial Least Squares Regression and Applications. National Defence Industry Press, Beijing (2000)
8. Helland, I.S.: On the Structure of Partial Least Squares Regression. Communications in Statistics—Simulation and Computation 17, 581–607 (1998)
9. Gao, H.X.: Statistical Analyses for Multiple Correlation Variables of Two Sets. Application of Statistics and Management 21(2), 58–64 (2002)

A Novel Numerical Random Model of Short Fiber Reinforced Foams

Yong Wu, Bo Wang, and Feiyu Han

College of Mathematics and Physics, Chongqing Institute of Technology,
Chongqing 400050, China
{wuyong,bowang,feih}@cqit.edu.cn

Abstract. Research on reinforced foams is getting heating, but the theoretic models of short fiber reinforced foams are still incomplete. This paper will discuss how a Random Sequential Adsorption Algorithm combines with a single tetrakaidecahedron of closed cell foams. As a result of research, a finite element numerical random analysis model will be set up. The model is used to compute the Young's modulus, strain and stress of the short fiber reinforced foams. The simulation results show that it is useful for advancing theories of the reinforced foams.

Keywords: Reinforced Foams; Short Fiber; Finite Element methods; Random Model.

1 Introduction

Presently, many researchers and the manufacturing industries are focusing on the research and application of short fibers reinforced foams. Hence, study on short fibers reinforced foams becomes a new topic of theoretical research and applications of composites. But the theoretical prediction of relative Young's Modulus of reinforced foams, is still a difficult issue. However, in recent yeas, there is still theoretic insufficiency in public literature on the reports of theories of the reinforced foams. Though satisfactory results of the Young's Modulus can be obtained by applying present finite element models in accordance with empirical modeling without reinforced short fibers with semi-empirical modeling with reinforced short fibers, these are only results of some idealized methods. Generally, the models are only used to some foams of low relative density. It might be able to make satisfactory results of prediction of the relative Young's Modulus to employ finite element methods to regular micro-structural cellular structure of the reinforced foams, but it is still difficult to obtain well prediction of Young's Modulus to some three-dimensional closed cell foams and two-dimensional closed cell foams structure with present finite element models or empirical modeling [1].

In this paper, a new numerical random model is being presented for enhancing theory of reinforced foams.

2 Numerical Model and Simulation

A representative volume element of short fibers reinforced foams, ie. a unit (cell), may be regarded as a tetrakaidecahedron, for instance, in Figure 1, with random short fibers insides of the cell walls. In the paper, the representative volume elements are

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 321 – 328, 2008.
© Springer-Verlag Berlin Heidelberg 2008

packed by repeating the units to computes the macro-properties of reinforced foams (shown in Figure 2). Here, a model of reinforced closed cell foam is created by a RSA (Random Sequential Adsorption) method. The basic principle is to randomly generate short fibers one by one in the cell walls of the tetrakaidecahedron above. If each new reinforced matter is put inside of the cell walls, and it overlaps with existing reinforced matter, a new random reinforced matter will be reproduced in accordance with place, length and orientation, and the new one will be reproduced if it is still overlapped. This process will be stopped if the number of reinforced matter of regeneration is over the previously fixed number [2,3].

The algorithm is as following:

At first, let us suppose that the number of short fibers in cylindrical shape which are continuously generated and rejected are k, if k is larger than the rejected total number K, then stop.

(1) Input parameters of probability distribution models to generate the random cylinder, and put the total number of cylindrical shapes N, or constitutive proportion parameters and K, the rejected total number, let n=1; k=0;

(2) Generate random cylindrical shape.

1). If n>N, or k>K, go to step (5), otherwise, go to next step;

2). To generate the base center, orientation and length of a cylindrical shape randomly by parameters distribution models;

3). If n=1, go to step (4), otherwise, go to step (3).

(3) If cylindrical shape is accepted, go to step (4), otherwise, set k=k+1, go to step (2);

(4) If a cylindrical shape placed there does not overlap any previously accepted cylinder, set n=n+1; if n<N, go to step (2), otherwise, go to step (5);

(5) Output.

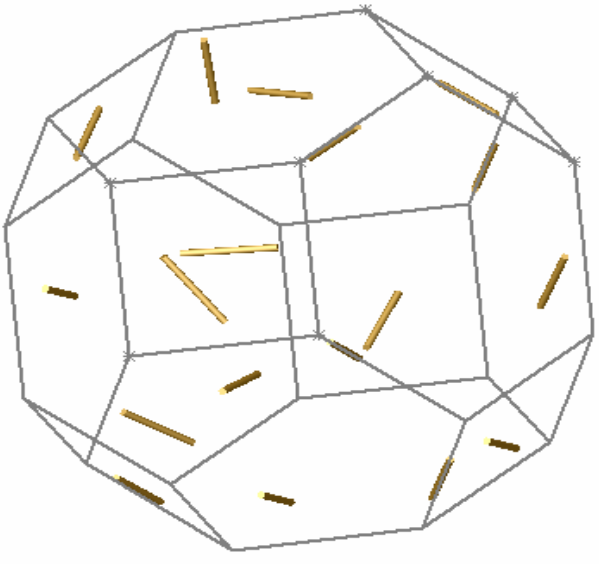


Fig. 1. Short Fibers of Uniform Distribution in Cell

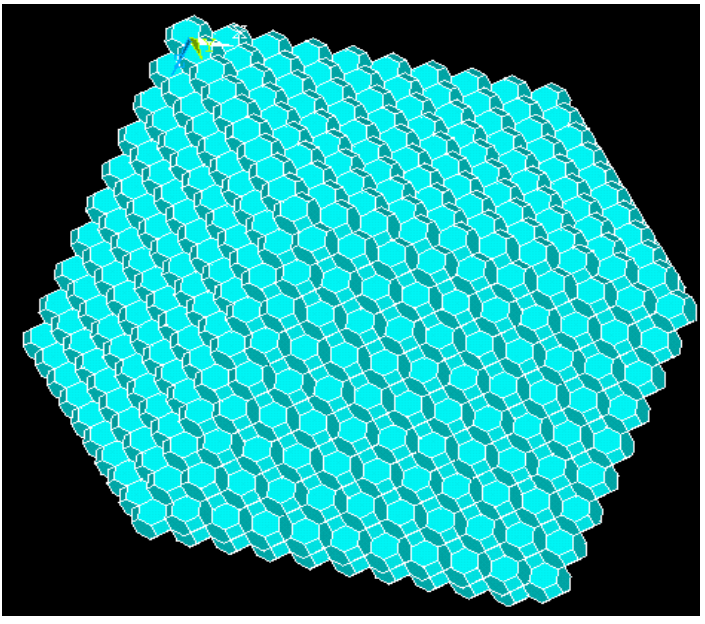


Fig. 2. Packing of 1729 Tetrakaidecahedrons

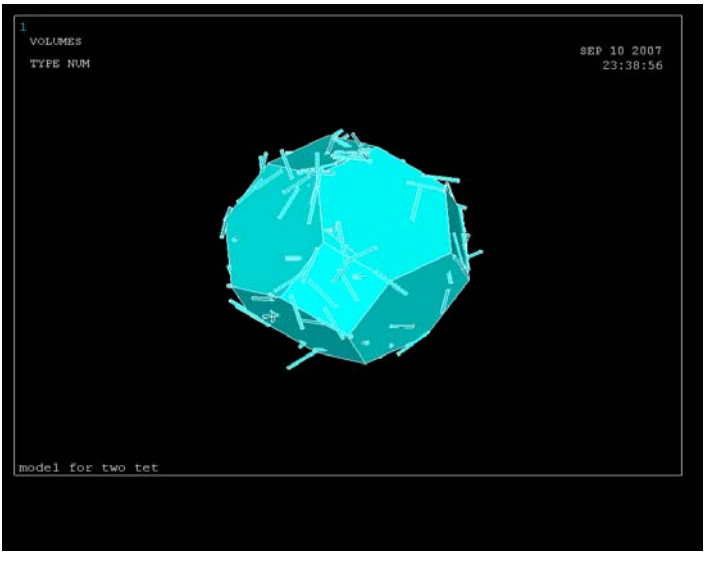


Fig. 3. Finite Element Model of Cell.

Figure 3 shows a simulation result of finite element method using above algorithm to put short fibers into the cell walls to volume fractions of 0.012 percent. Let short fibers be submitted to uniform distribution in range of the tetrakaidecahedron above,

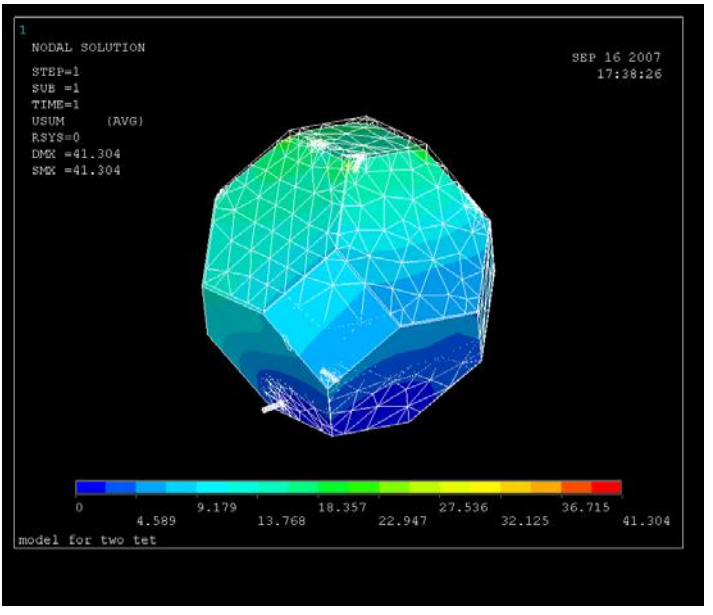


Fig. 4. Deformation

other distribution may be obtained by modifying the methods of generation of random numbers. Here, individual short fiber is permitted out of the cell walls, it conforms to reality[3].

Computing results of the micro-quantities using a unit model are as following (Fig. 4-Fig. 8), the mesh are auto-generating[4,5]:

Table 1. Basic Computational Parameters

Materials	Parameters of geometry	Density (g/cm^3)	Young's moduli (GPa)	Poisson's ratio
Matrix foams	2mm/1. 8mm	0. 9	1. 5	0. 33
Glass fibers of cylinder	L1mm, R0. 05	2. 6	72	0. 20

1). **Compressive Modulus Analysis:** the glass short fibers volume fractions are 0. 012 percent, the load is 1N along–Z orientation.

Through varying the quantity of short fibers in the cell, curve of Relative Young's Moduli of short fibers reinforce foam is obtained as shown in Fig. 6.

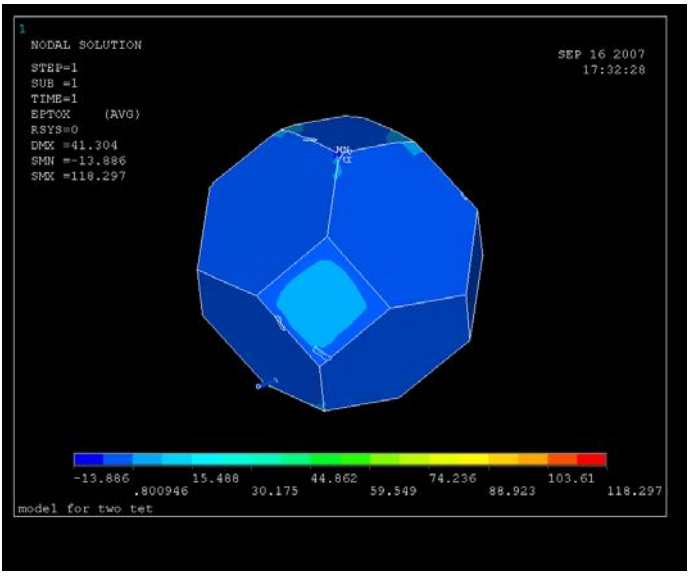


Fig. 5. Strain Distribution

Here, Relative Young's Moduli =Young's Moduli of short fibers reinforced foam/ Young's Moduli of matrix[1], Relative percentage=volume of short fibers/volume of reinforced foam.

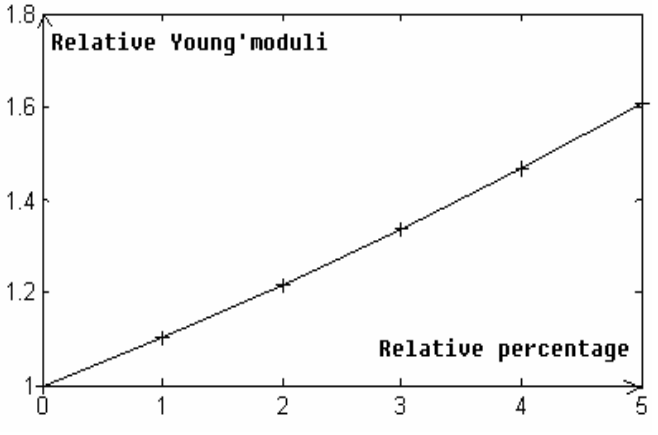


Fig. 6. Relative Young's Moduli

2). **Shear Modulus Analysis:** Volume fractions of glass short fibers are 0.012 percent. The load is 1N along-X orientation.

The research results above show that the random numerical model of closed cell foams is very effective, and of great value of applications [6].

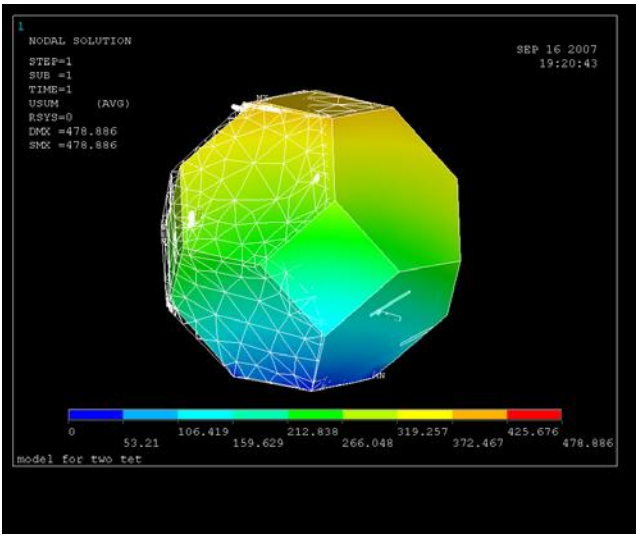


Fig. 7. Deformation

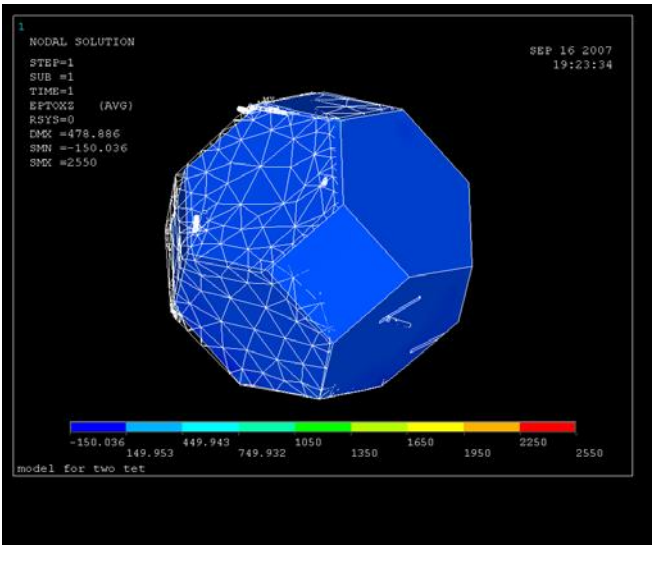


Fig. 8. Stress Distribution

3 Conclusions

In the paper, the numerical random model presented can be employed to predict relative Young's moduli of short fibers reinforced foams. This is a model of integration of two materials, in which non-continuous short fibers of random distribution exist in

matrix of continuous distribution. It can be employed to study relations of multiphase matter between the microstructure and mechanical macro-properties[6]. Hence, it is an innovation in the theory of reinforced foams. Computations of large scale will be conducted for gaining macro-properties of the reinforced foams in the future. The model to be advanced is not able to generate too high volume fractions of the reinforced matter. But when their volume fractions are not more than thirty percent, they are easy to be generated.

Acknowledgements. The work described in this paper was supported by grants from the National Natural Science Foundation of China (50573095), and the Natural Science Foundation of Chongqing in China (CSTC2007BB2411).

References

1. Lorna, J.G., Michael, F.A.: Cellular Solids, 2nd edn. Cambridge University Press, Cambridge (1999)
2. LLorca, J., Segurado, J.: Three-Dimensional Multiparticle Cell Simulations of Deformation and Damage in Sphere-Reinforced Composites. Materials Science and Engineering 365, 267–274 (2004)
3. Bohm, H.J., Eckschlager, A., Han, W.: Multi-Inclusion Unit Cell Models for Metal Matrix Composites with Randomly Oriented Discontinuous Reinforcements. Comp. Mater. Sci. 25(25), 42–53 (2002)
4. Burr, S.T., Vogel, G.D.: Material Model Development for Impact Analysis of Oriented Polypropylene Foam Structures. In: SAE 2001 World Congress, Detroit, Michigan (2001)
5. Roberts, P., Garboczi, E.J.: Elastic Moduli of Model Random Three-dimensional Closed-cell Cellular Solids. Acta Mater. 49, 189–197 (2001)
6. Roberts, A.P., Garboczi, E.J.: Elastic Properties of Model Random Three-dimensional Open-cell Solids. J. Mech. Physics Solids. 50, 33–55 (2002)

Appendix

1）Density, Relative Density, Volume and Relations of Geometry Size

(Omitted)

2) Basic Assumption

1. Foams materials composed of regular single tetrakaidecahedrons, sizes of all the units are uniform.
2. Short fibers are cylindrical and antisotropic, sizes of all fibers are similar. Its length is less than sides of the cell.
3. It is ideal for Fibers to combine with matrices.
4. Cross-section of cell walls of a tetrakaidecahedron is equilateral-triangle.
5. Fibers are uniformly distributed in the cell walls and boundary of tetrakaidecahedron. Fibers replace original place of matrices.

3）Basic Assumption of Parameters

1. Volume of the foams with or without fibers is V, density is ρ
2. Quantities of Relative Materials matrices is labeled as subscripts b, for instance, volume- V_b, density- ρ_b, etc.

3. Sidelength of tetrakaidecahedron is a, thickness of cell walls is c（c<<a）, sidetength of equilateral-triangle is b（b<<a）.
4. Length of fibers are l（l<a）, diameter is d（d<<l）, others are indicated by subscript f.
5. Percentage of fibers in foams is indicated by $x=V_f/V$.

4) Basic relations of Geometry

(Omitted)

Robust Wide Baseline Feature Point Matching Based on Scale Invariant Feature Descriptor

Sicong Yue, Qing Wang, and Rongchun Zhao

Northwestern Polytechnical University, Xi'an, P.R. China
typhoonadam@gmail.com

Abstract. This paper proposed a robust point matching method for wide baseline in order to achieve a large number of correct correspondences and high accuracy. To cope with large variations of scale and rotation, a feature descriptor, that is robust to scale and view-point, is added to the feature detection phase and it is included in the equations of the correspondence matrix that is central to the matching algorithm. Furthermore, the image window for normalized cross correlation is modified with adaptive scale and orientation. At the same time we remove from the matrix all the proximity information about the distance between points' locations which is the source of mismatches. Thus, the proposed algorithm is invariant to changes of scale, rotation, light and partially invariant to viewpoint. Experimental results show that the proposed algorithm can be used for large scene variations and provide evidence of better performance.

1 Introduction

A central problem of computer vision [1] is establishing correspondence between feature points in image sequence, for many complex tasks make use of point correspondence as an initial procedure, such as fundamental matrix estimation, image mosaic, motion analysis and tracking and 3D reconstruction from motion.

There have been many studies on this problem [2-8]. Classical approaches to point matching assume a short baseline and are usually based on correlation. It is well known that correlation based approaches suffer from changes of rotation and viewpoint. On this respect, an elegant approach falls in the family of spectral based methods by Scott and Longuet-Higgins [9].

The point matching method by Scott and Longuet-Higgins is based on a proximity matrix that depends on the distances between points' location. The method performs well on synthetic images, but it is sensitive to the noise and baseline changes in real images. More recently, Pilu [10] improved the algorithm by considering both proximity and similarity information in the correspondence matrix. The similarity is computed as the normalized correlation with the fixed window. Experimental shows that the performance of Pilu's method still drops in wide baseline.

We analysis the reason for this behavior and adopt a new feature and similarity measurement in the proposed algorithm. With such modification, a significant improvement is made on the performance of point matching. Several experiments on different types of image sequences (Fig.1) are presented.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 329 – 336, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The paper is organized as follows. Scale invariant feature descriptor is described in Section 2. Section 3 summarizes and explains the proposed robust wide baseline matching algorithm. Experimental results and conclusions are given in Section 4 and 5 respectively.

Fig. 1. Some examples of test image sequences. Top: 1st, 4th and 6th frame from boat sequence with zoom plus rotation. Middle: 1st, 4th and 6th frame from car sequence with illumination contrast. Bottom: 1st, 4th and 6th frame from graffiti sequence with viewpoint change.

2 Scale Invariant Feature Descriptor

To allow for efficient matching between wide baseline images, stable feature detection and descriptor across multiple views of a scene are required invariant to changes of rotation, scale, and illumination. Scale-space is an effective framework to handle objects at different scales. In Lowe's work [11], images are represented as a set of Scale Invariant Feature Transform (SIFT) features. Each SIFT feature represents a vector of local image measurements in a manner that is invariant to image translation, scaling, rotation and partially invariant to changes in illumination and affine deformations. The local and multi-scale nature of the SIFT feature makes them insensitive to noise, clutter and occlusion, and the nature makes them highly selective for matching to large amounts of features. From Mikolajczyk and Schmid [12] we get a comparative study on the effectiveness of the various feature descriptors proposed so far. Here

it is shown that the SIFT distance matcher (SDM) leads to best performance compared to other existing approaches.

The SIFT feature is localized directly on a scale space structure, searching local maxima of Difference of Gaussian (DoG) both on space and scale. At each such location, an orientation is selected at the peak of a histogram of local image gradient orientations. A feature vector is formed by measuring the local image gradients in a region around each location in coordinates relative to the location, scale and orientation of the feature. The gradient locations are further blurred to reduce sensitivity to noise and small local image deformations. In summary, the SIFT feature is stable across multiple views of a scene.

3 Robust Wide Baseline Matching Method

In this section we discuss the proposed robust wide baseline matching algorithm (RWM). Firstly, we will review some works of point matching based on spectral analysis. Spectral graph analysis aims at characterizing the global properties of a graph by using the eigenvalues and the eigenvectors of the graph adjacency matrix. Most of these contributions are based on so called proximity or affinity matrix. The matrix elements are weights that reflect the strength of a pair relation. Usually the proximity matrix is defined as:

$$G_{ij} = e^{-r_{ij}^2 / 2\sigma^2}, \tag{1}$$

where σ is a free parameter, and r_{ij} is the distance between points $p_i = I_1(x_i, y_i)$ and $q_j = I_2(x_j, y_j)$. $\mathbf{G}$ is a positive definite matrix and the element G_{ij} decreases monotonically from 1 to 0 with the increase of r_{ij}. Scott and Longuet-Higgins [9] were among the first to use spectral methods for image matching. They use the Euclidean distance between two points' location to measure the relation of two points. Scott's algorithm is based on the principle of proximity, that is, corresponding points must be the closest. More recently, Pilu [10] adds normalized cross correlation (NCC) of point's neighborhood in G_{ij} to quantify the similarity between points. The elements of $\mathbf{G}$ that model proximity and similarity can then be transformed as follow:

$$G_{ij} = \frac{C_{ij} + 1}{2} e^{-r_{ij}^2 / 2\sigma^2}, \tag{2}$$

where C_{ij} is NCC between two points' neighborhood. So G_{ij} has higher value when r_{ij} is smaller or C_{ij} is larger. Adding similarity constraint can eliminate rogue feature points, which shouldn't be similar to anything. But the performance of Pilu's method (PM) still drops with the growing of baseline.

We think the reason is that they adopt improper feature and similarity measurement in the matrix. On one hand, in the case of growing baseline, point localization is severely affected by changes of scale, rotation and affine transform. Thus, the distance between mismatch pair could be much closer than correct pair. On the other hand, the scale and orientation of the image window in NCC is not taken into account. Such

similarity measure can only handle the situation where there is translation or small changes of rotation and scale. So the performance will drop evidently.

Considering the problems analyzed above, we adopt more robust feature and similarity measure. From Mikolajczyk and Schmid [12], a comparative study of the performance of various feature showed that the SIFT distance matcher (SDM) is more robust than others with respect to changes of rotation, scale, viewpoint and affine transformation. The quality of the results decreases in the case of changes of illumination, for the SIFT feature is only invariant to linear brightness changes. Though SIFT has such a disadvantage, NCC with adaptive scale and orientation that is invariant to illumination changes can counterbalance the disadvantage.

Because the SIFT feature has such stable and distinctive character, it is more effective for image matching. We first use the distance between SIFT features ($\tilde{r}_{ij}$), which is similar to SDM, to measure the strength of a pair relation. And we discard all the proximity information about the distance between points' locations, which is the source of mismatches when the amount of scene variation increases.

And then, the similarity measure is modified by using NCC with adaptive scale and orientation (ASO-NCC). Both the size and the orientation of the correlation windows are determined according to the characteristic scale and orientation of feature points in contrast to the fixed size and fixed orientation in traditional NCC. So ASO-NCC is invariant to scale and rotation and it also can eliminate the affection of illumination. Let $p_i = I_1(x_i, y_i)$ be a point with scale s_1 and orientation θ_1 in one image I_1 and $q_j = I_2(x_j, y_j)$ be a point with scale s_2 and orientation θ_2 in the other image I_2. Without loss of generality, we can assume $s_1 > s_2$. W_1 and W_2 are two correlation windows of size $(2w+1)\times(2w+1)$ centered on each point with $w = \lambda s_2$, where λ is a constant ($\lambda = 5$ in this paper). Let $\theta = \theta_2 - \theta_1$ be the rotation angle and $r = s_1 / s_2$ be the scale change factor. W_1 and W_2 can then be represented as A_i and B_j :

$$\begin{cases} A_i^{uv} = I_1(x_i + ru\cos\theta + rv\sin\theta, y_i + rv\cos\theta - ru\sin\theta) \\ B_j^{uv} = I_2(x_j + u, y_j + v) \end{cases}, \tag{3}$$

where $u, v \in [-w, w]$. A_i^{uv} is the element of W_1 centered on p_i and B_j^{uv} is the element of W_2 centered on q_j. The similarity measure is then defined as:

$$\tilde{C}_{ij} = \frac{\sum_{u=-w}^{w}\sum_{v=-w}^{w}(A_i^{uv} - mean(A_i))(B_j^{uv} - mean(B_j))}{(2w+1)(2w+1)stdv(A_i)stdv(B_j)} \tag{4}$$

where $mean(A_i)$ is the average of A_i and $stdv(A)$ is the standard deviation of A_i.

With above modification, the correspondence matrix $\tilde{\mathbf{G}}$ is redefined as:

$$\tilde{G}_{ij} = \frac{\tilde{C}_{ij} + 1}{2} e^{-\tilde{r}_{ij}^2 / 2\sigma^2}. \tag{5}$$

The next step is to perform the singular value decomposition of $\tilde{\mathbf{G}}$:

$$\tilde{\mathbf{G}} = \mathbf{U}\mathbf{D}\mathbf{V}^T . \tag{6}$$

Then compute a new pairing matrix $\mathbf{P}$ by converting diagonal matrix $\mathbf{D}$ to a diagonal matrix $\mathbf{E}$ where each D_{ii} is replaced with 1.

$$\mathbf{P} = \mathbf{U}\mathbf{E}\mathbf{V}^T . \tag{7}$$

$\mathbf{P}$ has the same shape as $\tilde{\mathbf{G}}$ and it has the interesting property of sort of amplifying good pairs and attenuating bad ones. Two points p_i and q_j are pairing up if P_{ij} is both the greatest element in its row i and the greatest in its column j. No threshold is required for choosing good pair.

4 Experimental Results

To validate the proposed algorithm, experiments have been carried out on several image sequences against Pilu's method (PM) and the SIFT distance matcher (SDM) mentioned in section 3. The test image sequences include different geometric and photometric transformation, such as large rotation and scale variance, viewpoint change and illumination contrast. The variance grows larger with the increasing of the frame index. Some examples are shown in Fig.1. All the image sequences are from INRIA (http://lear.inrialpes.fr/people/Mikolajczyk/Database/index.html).

In all case we fixed the first frame (index '0') of each sequence as the reference image and made a match between the other frame and it. A robust estimation of the homography transformation between two images based on RANSAC enables the selection of correct matches, as used in [12].

The results in Fig.2 are related to boat sequence of 6 frames where there are different degree of rotation and scale variance. Fig.3 shows the results in car sequence of 6 frames with different light contrast. Table 1 gives the results of experiment on graffiti sequence with viewpoint changes. Because the results of last 2 frames in graffiti sequence are too bad, only the matching results of 2nd, 3rd and 4th frames are given in Table 1. From results, it can be found that the accuracy of RWM and SDM are almost on a par for large rotation, scale and viewpoint changes (Fig.2, Table 1) while that of PM is much lower; for large illumination contrast (Fig.3) the accuracy of RWM and PM are almost the same but SDM drops evidently.

Apart from the accuracy, the correct match number is another more important the figure of merit for wide baseline matching. In all experiments, the proposed algorithm RWM returns the largest number of correct matches. Especially for the last few frames, RWM can still obtain desirable large number of correct matches that is several times of PM and SDM. In the last frame of boat sequence (Fig.2) and graffiti sequence (Table 1), SDM and PM only return 4-6 correct matches, while RWM gets 25-50 correct pairs. For boat sequence and graffiti sequence, SDM performs better than PM, that is because the SIFT feature is more robust to scale, rotation and viewpoint changes than NCC. But PM performs better than SDM in car sequence with

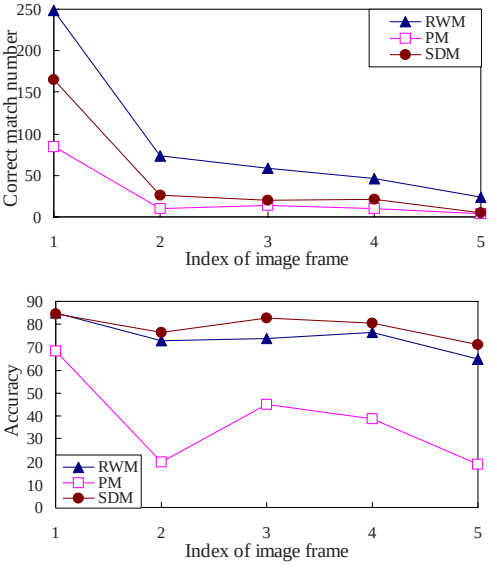


Fig. 2. Curves of correct match number and accuracy: comparison results for boat sequence with scale and rotation changes. Top: number of correct matches detected by three methods. Bottom: accuracy for three matching methods.

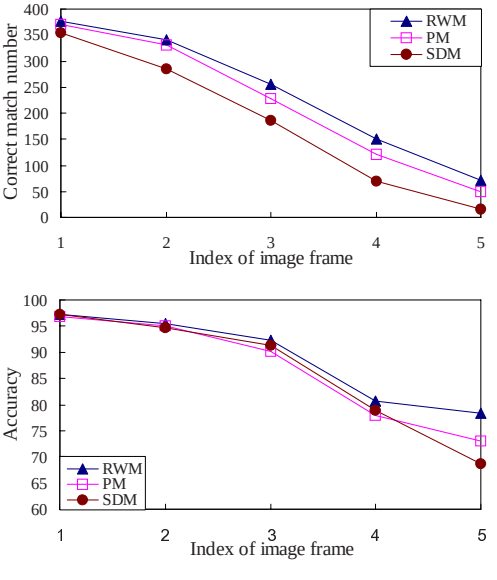


Fig. 3. Curves of correct match number and accuracy: comparison results of car sequence with illumination contrast. Top: number of correct matches detected by three methods. Bottom: accuracy for three matching methods.

Table 1. Correct matches and accuracy of graffiti sequence. For the results of last 2 frames are totally failed, only the results of 2nd, 3rd and 4th frames are given.

Method	RWM	PM	SDM	RWM	PM	SDM	RWM	PM	SDM
Index of frame		1			2			3	
Correct matches	679	406	575	302	64	187	45	5	5
Accuracy	0.92	0.88	0.95	0.76	0.37	0.65	0.43	0.08	0.45

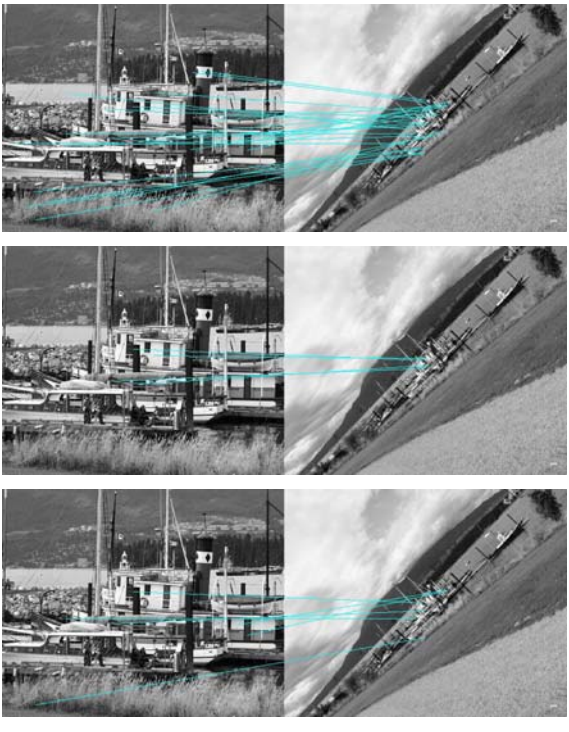


Fig. 4. Example of matching results for three methods. Top: RWM. Middle: PM. Bottom: SDM. The pair of images is the 1st and 6th frame from boat sequence with large scale and rotation changes.

illumination change, which shows NCC is more robust to general light changes than SIFT features. Some examples of matching results from above experiments are shown in Fig.4.

5 Conclusions

In this paper we presented a new robust matching method based on spectral graph analysis for wide baseline. By discarding the unreliable spatial distance between

points' location and adopting SIFT feature and ASO-NCC, the correspondence matrix enhances the similarity constrain on the optimal matching pair and it has the advantage of reducing the risk of mismatches with respect to the previous proximity matrix. Satisfactory matching results have been achieved when large baseline changes occur in the scene. Experimental evidence shows that the proposed algorithm returns better performance for scale changes, image plane rotation, large illumination contrast and viewpoint transform as compared to Pilu's matching methods and the SIFT distance matcher. A fly in the ointment is that the proposed algorithm is only partially invariant to affine transformation caused by viewpoint, although the other two methods are even worse. Therefore the use of affine invariant feature in correspondence matrix is the next step for the further study. It might then be combined with SIFT features to make this algorithm more robust and reliable.

Acknowledgments. The work in this paper was supported by National Hi-Tech Development Programs of China under grant No. 2007AA01Z314 and New Century Excellent Talents in University (NCET-06-0882), P. R. China.

References

1. Marr, D.: Early Processing of Visual Information. Philosophical Transactions of the Royal Society of London (B) 275, 483–524 (1976)
2. Zitov, B., Flusser, J.: Image Registration Methods: A Survey. Image and Vision Computing 21, 977–1000 (2003)
3. Xiaoye, L., Manduchi, R.: Wide Baseline Feature Matching Using the Cross Epipolar Ordering Constraint. In: Proceedings of IEEE Conference on CVPR, pp. 16–23. IEEE Press, New York (2004)
4. Pritchett, P., Zisserman, A.: Wide Baseline Stereo Matching. In: Proceedings of the International Conference on Computer Vision, pp. 754–760. IEEE Press, New York (1998)
5. Zhang, Z., Deriche, R., Faugeras, O., Luong, Q.T.: A Robust Technique for Matching two Uncalibrated Images through the Recovery of the Unknown Epipolar Geometry. Artificial Intelligence 78, 87–119 (1995)
6. Ferrari, V., Tuytelaars, T., Gool, L.V.: Wide-baseline multiple-view correspondences. In: Proceedings of IEEE Conference on CVPR, pp. 718–725. IEEE Press, New York (2003)
7. Carcassoni, M., Hancock, E.: Spectral correspondence for point pattern matching. Pattern Recognition 36, 193–204 (2003)
8. Sclaroff, S., Pentland, A.: Modal Matching for Correspondence and Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 17, 545–561 (1995)
9. Scott, G., Longuet-Higgins, H.: An algorithm for associating the features of two patterns. In: Proceedings of the Royal Society of London, pp. 21–26 (1991)
10. Pilu, M.: A Direct Method for Stereo Correspondence Based on Singular Value Decomposition. In: Proc. of IEEE Conference on CVPR, pp. 261–266. IEEE Press, New York (1997)
11. Lowe, D.: Distinctive Image Features from Scale-Invariant Key-Points. International Journal of Computer Vision 60, 91–110 (2004)
12. Mikolajczyk, K., Schmid, C.: A Performance Evaluation of Local Descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence 27, 1615–1630 (2005)

Detection and Recognition of Scoreboard for Baseball Videos

Chaur-Heh Hsieh[1], Chin-Pan Huang[1], and Mao-Hsiung Hung[2]

[1] Dept. of Computer and Communication Engineering, Ming Chuan University
Taoyuan, Taiwan
hsiehch@mail.mcu.edu.tw
[2] Dept. of Information Engineering, I-Shou University
Kaohsiung, Taiwan

Abstract. This paper presents an effective and efficient detection and recognition of scoreboard caption method for baseball videos. The method first identifies the scoreboard type using template matching and then extracts the caption region of each type. Next it recognizes the extracted caption utilizing a novel digit recognition scheme which is constructed by a simple neural network classifier. It results in a much simpler method with significantly higher recognition rate over that of the universal OCR scheme. Experimental results demonstrate the effectiveness of the proposed method and indicate that it identifies twelve scoreboard types correctly and recognizes scoreboard caption over 98%.

Keywords: Detection and recognition of scoreboard caption, OCR, neural network classifier.

1 Introduction

Over the past decade, the amount of multimedia information has grown explosively. This trend necessities the development of efficient content-based indexing and retrieval techniques. Recently, automatic sports video analysis has received increasing attention, because sport video appeals to large audiences. The processing of sports video generates various valuable applications such as highlighting, summarization, indexing/retrieval, athlete's training and entertainment. The application systems make the delivery and searching of video contents more effective and efficient. For example, highlight or summarization makes it easy to deliver sports video over even narrow band networks, such as the Internet and hand-held wireless, since the valuable semantics generally occupy only a small portion of the whole content. Due to the automatic indexing of sport contents, users can retrieve their preferred particular clips of sport videos such as hits of baseball or goals of soccer.

The major challenge of sport video analysis is how to bridge the gap between low-level features and high-level semantic. In broadcast sport videos, superimposed scoreboard is used to display game status such as team names, score, time, etc., to increase the audience's understanding of the game program. It provides a good and reliable cue for high-level semantics analysis.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 337 – 346, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Recently, several general algorithms of caption recognition have been published. Tang et al [1] proposed a universal caption detection and recognition method based on a fuzzy-clustering neural network technique. Yan et. al. [2] presented a video logo detection scheme to manage overlapped types and blended types. Zhang et. al. [3] proposed general and domain-specific techniques. They first present a general algorithm to detect and locate caption, and then they employ domain model of specific sports, e.g. baseball and basketball, in the word recognition to improve its recognition rate from 78% to 87%.

However, these general algorithms are too complicated and their recognition performance is still not high enough for the scoreboard caption in broadcast sport videos. Fig. 1 shows a typical baseball scoreboard which contains game status including inning number, ball count (number of strike (S) and ball (B)), and number of outs (O), scores for two teams, base occupation status, and indication of offensive side. For a particular scoreboard style, the field boxes are fixed or only slightly changed; namely, the font type and relative location of each field are kept the same over the whole video. Thus, if the scoreboard style can be identified, we can use the prior knowledge of the fixed field boxes to achieve the benefits of simplicity and low computation in implementation.

This paper presents an effective and efficient method to detect and recognize the scoreboard caption in baseball videos. In the initial phase, the scoreboard type identification is first performed using template matching. According to the identified scoreboard type, the location of each field in a scoreboard can be easily obtained based on prior established knowledge. The digit of each field box is then identified using a novel digit recognition scheme which is constructed by a simple neural network classifier, derived by using 22 features extracted from the image contour. Experimental results indicate that the proposed technique achieves very high recognition rate with relatively low complexity in implementation.

Fig. 1. A typical style of scoreboard (base occupation: 1st and 2nd bases; score: 7 vs 10; inning: 9th-top; O:1, B:2, S:2)

2 Proposed Scoreboard Detection and Recognition Method

The proposed method consists of two processes, location process and recognition process, as shown in Fig. 2. In the location process, it first finds out the position of a scoreboard, and then determines scoreboard type and the location of each field of the scoreboard. This process only performs for the initial part of a test video, called initial clip. In the recognition process, it identifies the data of each field of a scoreboard for a whole video. The details are described in the following.

2.1 Location Process

Scoreboard Type Identification. The flow chart of the scoreboard type identification is shown in Fig. 3. The purpose of this stage is to determine which

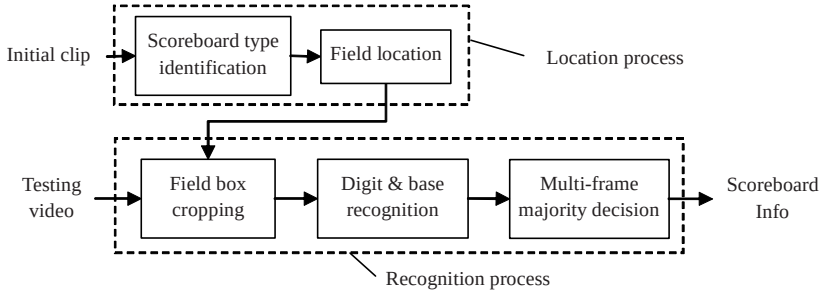


Fig. 2. Flow chart of overall scoreboard detection and recognition

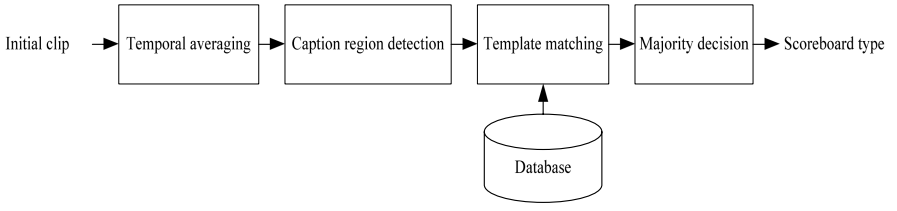


Fig. 3. Flow chart of scoreboard type identification

scoreboard type is embedded in the testing video. The identification procedure contains four steps as follows:

(1) Temporal averaging calculates the average intensity through a few (say 300) consecutive frames of an initial clip of a video. The procedure is similar to that of our previous work [4]. Superimposed caption is attached at fixed location in the video, and generally the caption always exists in all video frames. Therefore, after temporal averaging, the superimposed caption still remains, whereas the other video contents would disappear.

(2) A simple binarization detects the caption region from the average image. Given the average image $A(x,y)$, the binary image $B(x,y)$ is obtained by Eq.(1)

$$B(x,y)=\begin{cases}0(\text{non-caption pixel}),\\ \text{if } (\mu-k\cdot\sigma)\le A(x,y)\le(\mu+k\cdot\sigma)\\ 1(\text{caption pixel}),\ \text{otherwise.}\end{cases} \tag{1}$$

where μ and σ respectively represent the mean and the standard deviation of the average image, and k is assigned 1.75 experimentally.

(3) The caption region detection obtains some scoreboard candidates. We match these candidates with the scoreboard templates stored in database, which is established as follows. To achieve the generality of our scheme, we have collected twelve kinds of scoreboards from different broadcast channels including ESPN, FOX, YES, STARSPORT, and VIDEOLAND SPORTS and generated their blurred versions of the these scoreboard images as templates, instead of the original images. The reason for using the blurred images as templates is as follows. Since the contents of a scoreboard changes over different games and times, e.g. team names, score, and so on. The

content change decreases the accuracy of template matching. The blurred image attacks the problem because the content details are smoothed out. Fig. 4 shows two of twelve templates which are stored in the database. The shape of template is obtained manually by setting scoreboard pixels as 1 and the remaining as 0. The blurred version of the scoreboard image is regarded as texture image. The shape and texture are used for the identification of scoreboard types.

Given K templates in the database, denoted as T_k, k=1, 2, …, K. The caption matching evaluates shape and texture similarity between a test video and the templates. The shape and texture (average image) of the detected caption region of the test video are respectively denoted as α and F. The template matching consists of two steps: calculate shape similarity first, and texture MAD (mean absolute difference) later. The shape similarity between an unknown region X and the template T_k is defined as their intersection area by

$$\text{Shape similarity}(X, T_k) = \frac{\text{Area}\left(\alpha_X \cap \alpha_{T_k}\right)}{\text{Area}\left(\alpha_{T_k}\right)} \tag{2}$$

Those templates (T_k'), whose shape similarities against X are larger than a predefined threshold, continue the next step. In the second step, MAD of X and T_k' is calculated by

$$\text{MAD}(X, T_k') = \sum_{(x,y)\in\alpha_X\cap\alpha_{T_k'}} \left|F_X(x,y) - F_{T_k'}(x,y)\right| \Big/ \text{Area}\left(\alpha_X \cap \alpha_{T_k'}\right) \tag{3}$$

The template which generates the smallest MAD determines the scoreboard type of the test video.

(4) To achieve robust identification of scoreboard types, we further apply majority voting after template matching. An initial video clip is partitioned into several short clips of 300-frame long. Every clip can determine a vote for one of templates. The testing video is classified into the template which gets majority votes.

Our experiment result indicates 100% of identification rate is obtained by using both shape and texture matching. However, 75% of accuracy can be achieved if only shape matching is employed.

Fig. 4. Original caption and its blurred version

Field Location. A scoreboard contains several game status fields including team names, base, score, inning, out#, ball# and strike#. The positions of all fields for a particular scoreboard type are fixed, so the field position information can be obtained by manual labeling. Fig. 4 illustrates field locations of a typical scoreboard type. The locations of the important fields (marked by green rectangle) of all scoreboard types are pre-stored in a database. Once

the type of a scoreboard is determined , its corresponding field location information can be obtained by checking the database, and then the results are sent to the field box cropping of the recognition process.

2.2 Digit Recognition

In the recognition phase, the image of each field is cropped out according to the field location information obtained from the location process mentioned above. These images including base status and digit data are then identified. The base status can be easily identified using color information since it changes from one color to another when the base is occupied. The recognition of digits of 0 to 9 includes three stages, preprocessing, feature extraction and neural network classifier, as shown in Fig. 5. Firstly, the preprocessing binarizes and partitions digit boxes from the original frame. Then, feature extraction calculates a set of feature vector representing the digit inside the cropped box. Finally, a simple neural network classifier, obtained by a prior training process, identifies the input digit. The detailed is described in the following.

Preprocessing. It executes the following steps in this stage.

Step 1: size enlargement
Step 2: binarization using three membership functions for RGB components
Step 3: morphological erosion
Step 4: size normalization based on the minimum bounding box of the binary image
Step 5: extraction of outside contour and inside contour of a digit

For the binarization, we define three membership functions $\mu_R(r)$, $\mu_G(g)$ and $\mu_B(b)$, for R, G, and B components, respectively, using the domain knowledge of the scoreboard styles. Our investigation indicates that the image features of scoreboards vary with different scoreboard styles. Thus, we define several sets of membership functions for various styles, as shown in Fig. 6.

The binary image can be obtained by Eq.(4).

$$B(r,g,b) = \begin{cases} 1, \text{ if } \dfrac{\mu_R(r) + \mu_G(g) + \mu_B(b)}{3} > TH \\ 0, \text{ otherwise.} \end{cases} \tag{4}$$

TH is assigned 0.65 experimentally. After morphological erosion and size normalization, we extract the inside and outside contours of the digit image for the following processing, as shown in Fig. 7.

Feature Extraction. To extract the orientation features of the contour, we design eight types of binary patterns, as shown in Table 1. The first two types are used to detect horizontal and vertical features of a digit, respectively. The next three types are used to detect slash features with different orientations. In addition, three types of backslash features are defined as well. To explain the feature extraction process, take the first pattern (horizontal pattern) as an example. The horizontal pattern contains two white (logical 1) pixels, which is regarded as a 1x2 mask (window). Move the mask over the input digit image in a raster scan manner (i.e. shift pixel-by-pixel from left to right and top to bottom), and count the total number of matches (called match

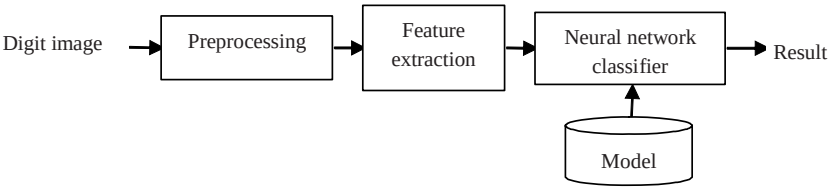


Fig. 5. Flow chart of digit recognition

frequency hereafter). If the input digit data in the window at a particularly time (window is moving) are completely equal to the horizontal pattern, then it is a match; otherwise a mismatch. The feature extraction process for the other patterns is the same.

The count of contour pixel inside a particular zone is also defined as a feature. Four zone partition types including horizontal, vertical, slash and backslash are shown in Table 2. Each partition type contains two zones, so it generates two feature values.

For the outside contour, all types of orientation features, denoted as v_1, v_2,…, v_8, are used. For the inside contour, only four types of orientation features including horizontal, vertical, slash I and backslash I, denoted as v_9, v_{10},…, v_{12}, are considered. For the outside contour, all types of zoning features, represented as v_{13}, v_{14},…, v_{20}, are used. For the inside contour, only the horizontal zone is considered and its features are denoted as v_{21} and v_{22}. In summary, a 22-D feature vector from a digit image is extracted. For the digits "1", "2", "3", "5" and "7", there are no inside contour, so zero is padded for the corresponding dimensions including (v_9, v_{10},…, v_{12}) and (v_{21}, v_{22}).

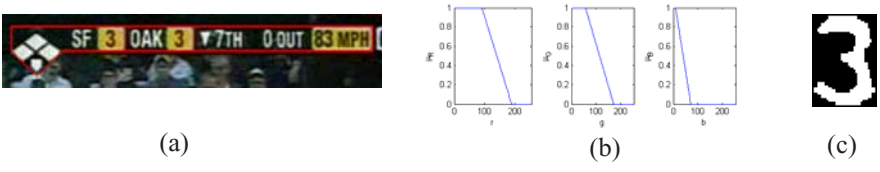


Fig. 6. Binarization for typical scoreboard style: (a) original images of two scoreboard styles, (b) membership functions for binarization, (c) binarized images of score fields after erosion

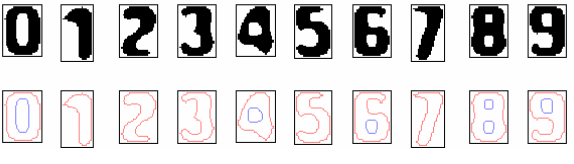


Fig. 7. Outside and inside contours of 10 digits

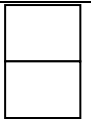
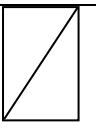
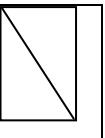
For the outside contour, all types of orientation features, denoted as v_1, v_2,…, v_8, are used. For the inside contour, only four types of orientation features including

horizontal, vertical, slash I and backslash I, denoted as v_9, v_{10},…, v_{12}, are considered. For the outside contour, all types of zoning features, represented as v_{13}, v_{14},…, v_{20}, are used. For the inside contour, only the horizontal zone is considered and its features are denoted as v_{21} and v_{22}. In summary, a 22-D feature vector from a digit image is extracted. For the digits "1", "2", "3", "5" and "7", there are no inside contour, so zero is padded for the corresponding dimensions including (v_9, v_{10},…, v_{12}) and (v_{21}, v_{22}).

Table 1. Orientation patterns of eight types

$\begin{bmatrix}1 & 1\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$	$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
Horizontal	Vertical	Slash I	Backslash I
$\begin{bmatrix}0 & 1\\0 & 1\\1 & 0\end{bmatrix}$ or $\begin{bmatrix}0 & 1\\1 & 0\\1 & 0\end{bmatrix}$	$\begin{bmatrix}1 & 0\\0 & 1\\0 & 1\end{bmatrix}$ or $\begin{bmatrix}1 & 0\\1 & 0\\0 & 1\end{bmatrix}$	$\begin{bmatrix}0 & 0 & 1\\1 & 1 & 0\end{bmatrix}$ or $\begin{bmatrix}0 & 1 & 1\\1 & 0 & 0\end{bmatrix}$	$\begin{bmatrix}1 & 0 & 0\\0 & 1 & 1\end{bmatrix}$ or $\begin{bmatrix}1 & 1 & 0\\0 & 0 & 1\end{bmatrix}$
Slash II	Backslash II	Slash III	Backslash III

Table 2. Zoning partition of four types

			
Horizontal	Vertical	Slash	Backslash

Neural Network Classifier. A neural network of single layer of ten neurons is employed to classify ten digits, as shown in Fig. 8. Log-Sigmoid of transfer function and Widrow-Hoff learning rule are employed [5]. In training stage, the digits of twelve caption styles (i.e. twelve digit fonts) were collected, and their feature vectors using the scheme mentioned above were extracted. Then, applying the feature vectors into learning until mean square error (MSE) has been converged below an acceptable value, or a preset maximum number of iterations have been reached. In testing stage, the neuron which generates maximum output value determines which digit is inputted to the network. The experimental results indicate the single-layer network achieves very good performance.

Widrow-Hoff learning rule is used to compute weights and bias of the network iteratively. At k-th iteration, $\mathbf{p}(k)$ is an input to the network, and $\mathbf{t}(k)$ is the corresponding target output. Then, the output is

$$\mathbf{a}(k) = f\left(\mathbf{W}(k)^T \mathbf{p}(k) + \mathbf{b}(k)\right) \tag{5}$$

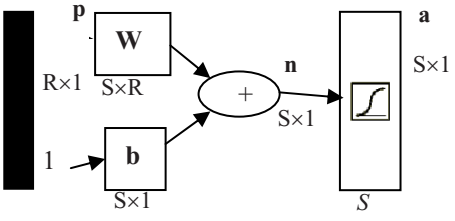
The error vector at the *k*-th iteration is

$$\mathbf{e}^2(k) = \begin{bmatrix} e_1^2(k) \\ \vdots \\ e_S^2(k) \end{bmatrix} = \begin{bmatrix} \left({}_1\mathbf{t}(k) - {}_1\mathbf{a}(k)\right)^2 \\ \vdots \\ \left({}_S\mathbf{t}(k) - {}_S\mathbf{a}(k)\right)^2 \end{bmatrix} \tag{6}$$

where ${}_i\mathbf{t}(k)$ and ${}_i\mathbf{a}(k)$ mean *i*th element of the vectors. The LMS algorithm adjusts weights and bias to minimize the mean square error at (*k*+1)-th iteration as

$$\begin{cases} \mathbf{W}(k+1) = \mathbf{W}(k) + 2\alpha\,\mathbf{e}(k)\mathbf{p}^T(k) \\ \mathbf{b}(k+1) = \mathbf{b}(k) + 2\alpha\,\mathbf{e}(k) \end{cases} \tag{7}$$

where α is a constant learning rate.



PS: R=22, S=10

Fig. 8. Neural network of single layer of ten neurons

2.3 Multi-frame Majority Decision

After digit and base recognition, one result from a single frame is obtained. Because the data of each field may change after occurrence of a new event, the data of the same field generally keep the same for a relative long time. This characteristic can be employed to further improve the recognition rate. In this work, we use majority voting for several consecutive frames to correct the recognition errors of few frames. It is noted that a vote is made from the results of the consecutive frames belonging to the same shot, e.g. pitch shot.

3 Experimental Results

In the experiments, baseball videos from seven baseball games of American League Champion Series (ALCS) of 2004 MLB playoffs are used. Those videos are denoted as ALCS 1~7, where ALCS 1, 2, 6 and 7 were taken place in Yankee Stadium (home of New York Yankees) and ALCS 3, 4 and 5 in Frenway Stadium (home of Boston Red Sox). All the videos are compressed in MPEG-1 format with frame size of 352x240 and frame rate of 30 fps. Experimental results are summarized in the following.

Table 3. Results of scoreboard recognition of proposed scheme

Games	Base	Scores of two teams	Top/bottom inning	# of out
ALCS 1	99% (78/79)	100% (158/158)	100% (79/79)	100% (79/79)
ALCS 2	100% (71/71)	100% (142/142)	100% (71/71)	100% (71/71)
ALCS 3	97% (95/98)	99% (195/196)	100% (98/98)	100% (98/98)
ALCS 4	100% (107/107)	99% (212/214)	99% (106/107)	97% (104/107)
ALCS 5	94% (118/126)	99% (250/252)	98% (123/126)	95% (120/126)
ALCS 6	99% (74/75)	94% (148/150)	99% (74/75)	99% (74/75)
ALCS 7	98% (82/84)	99% (166/168)	100% (84/84)	98% (83/84)
Average accuracy rate	98% (625/640)	99% (1271/1280)	98% (635/640)	98% (629/640)

$$\text{PS: accuracy rate} = \frac{\#\text{of correct detected}}{\#\text{of ground truth}} \times 100\%$$

Five items of information on the scoreboard including occupied base, scores, top/bottom inning and number of out are planned to identify. The goal is achieved by using the scheme described in Section 2. The results of scoreboard recognition are listed in Table 3. It indicates that average recognition rate achieved is over 98%. In addition, the performance is rather stable for all games.

The commercial OCR SDK, designed for all of alphabets, digits and symbols, of Lab Asprise [6] is utilized in experiment. It is applied to recognize score and out# captions in videos, and its results are compared with our scheme in Table 4. It indicates our scheme performs significantly better than OCR does. In addition, our scheme is much simpler than OCR in implementation.

Table 4. Comparison of proposed method and OCR

Games	Scores of two teams		# of out	
	Proposed	OCR	Proposed	OCR
ALCS 1	100% (158/158)	93% (147/158)	100% (79/79)	86% (68/79)
ALCS 2	100% (142/142)	99% (141/142)	100% (71/71)	75% (58/71)
ALCS 3	99% (195/196)	94% (185/196)	100% (98/98)	97% (95/98)
ALCS 4	99% (212/214)	77% (165/214)	97%(104/107)	81% (87/107)
ALCS 5	99% (250/252)	81% (204/252)	95% (120/126)	81% (102/126)
ALCS 6	94% (148/150)	87% (131/150)	99% (74/75)	73% (55/75)
ALCS 7	99% (166/168)	69% (116/168)	98% (83/84)	75% (63/84)
Average accuracy rate	99% (1271/1280)	85% (1089/1280)	98% (629/640)	83% (528/640)

4 Conclusions

An effective and efficient method for detecting and recognizing baseball scoreboard caption in broadcast videos has been presented in this paper. The method consists of three novel schemes including scoreboard type identification, field box location and

digit recognition. It achieves more than 98% of recognition rate for the baseball scoreboard, which is significantly better than the conventional OCR. In addition, the appropriate usage of the domain model makes the system much simpler than OCR in implementation. The extension to other sport videos is currently investigated.

Acknowledgments. This work was supported by National Science Counsel Granted NSC 95-2221-E-214-051.

References

1. Tang, X., Gao, X., Liu, J., Zhang, H.Z.: A Spatial-temporal Approach for Video Caption Detection and Recognition. IEEE Trans.on Neural Networks 13, 961–971 (2002)
2. Yan, W.Q., Wang, J., Kankanhalli, M.S.: Automatic Video Logo Detection and Removal. Multimedia System 10, 379–391 (2005)
3. Zhang, D., Rajendran, R.K., Chang, S.F.: General and Domain-Specific Techniques for Detecting and Recognizing Superimposed Text in Video. In: Int. Conf. on Image Processing, pp. 22–25 (2002)
4. Su, Y.M., Hsieh, C.H.: A Novel Model-based Segmentation Approach to Extract Caption Contents on Sports Videos. In: ICME 2006, pp. 1829–1832 (2006)
5. Widrow, B., Hoff, M.E.: Western Electric Show and Convention Record, IRE, part 4, pp. 95-104 (1960)
6. Asprise OCR, `http://asprise.com/`

A Heuristic Approach to Caption Enhancement for Effective Video OCR

Lei Xie[1,2] and Xi Tan[1,3]

[1] School of Computer Science, Northwestern Polytechnical University, Xi'an, China
[2] Human-Computer Communications Laboratory
The Chinese University of Hong Kong, Hong Kong SAR
[3] Tsinghua-CUHK Joint Research Center for Media Sciences, Technologies and Systems, Shenzhen, China
lxie@nwpu.edu.cn

Abstract. We present a heuristic approach to enhancing speech synchronized captions for video OCR, as a pre-process for subsequent tasks of multimedia indexing, segmentation and retrieval. We use a bi-search based caption transition detection method to improve efficiency, which adopts a simple heuristics that the same caption content usually lasts for a period for stable viewing. We propose a combination of color mask, changing mask and region mask to perform caption enhancement based on the discriminative characteristics of captions and backgrounds. Elaborate enhancement on individual characters is further used to remove small background residues. OCR experiments show that our caption enhancement approach brings a high character accuracy of 89.24%.

Keywords: Video OCR, caption transition detection, caption enhancement, video retrieval, multimedia retrieval.

1 Introduction

With the proliferation of videos from TV and the Web, extraction of superimposed captions in videos (i.e. Video OCR) has drawn much interest for multimedia segmentation, categorization and retrieval [1,2,3]. Video captions contain rich content information such as names, locations, dates and time. Speech-synchronized captions in news archives are valuable for efficient news indexing and retrieval. However, extracting caption texts in videos is not a trivial task. Mature OCR technologies for document images cannot be directly used to recognize video captions since caption characters are often embedded in complex and dynamic instead of clear and static backgrounds as in document images. Captions usually hold small areas with low resolution characters and lossy video compression further degrades the quality. Video segmentation and categorization tasks desire the need of extracting the timing information of captions.

Approaches in Video OCR have mainly focused on caption/text detection and enhancement [4,5,6]. Caption detection locates the caption regions and finding out caption (dis)appearance frames (i.e. caption transition detection). Upon detection of captions, image enhancement is adopted to enlarge the discrimination between caption characters and backgrounds. Tang *et. al.* have proposed a

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 347–355, 2008.
© Springer-Verlag Berlin Heidelberg 2008

spatial-temporal approach for Chinese caption detection and enhancement [4], where fuzzy clustering neural network (FCNN) is introduced for caption transition detection; multi-frame integration methods (e.g. minimal pixel search and frame averaging) and morphological operations are adopted for caption enhancement. FCNN-based caption transition detection is effective but it induces frame-by-frame feature extraction and classification that result in a low efficiency. Minimal pixel search for caption enhancement may not necessarily lead to a high caption-background contrast when backgrounds and captions share a similar or nearly the same color, which degrades the OCR performance seriously.

In this paper, we present a heuristic approach to enhancing speech synchronized captions for effective Video OCR as a pre-process for subsequent tasks of news segmentation and retrieval. We use a bi-search based caption transition detection method that improves efficiency using a simple heuristics that the same caption text usually lasts for a bunch of frames. We propose three mask operations (i.e. color mask, changing mask and region mask) to perform caption enhancement based on the discriminative characteristics of captions and backgrounds. Elaborate enhancement on individual characters is further adopted to remove small background residues. Our approach outperforms Tang's approach [4] tested on the videos from the same source.

2 System Overview

The block diagram of our system is shown in Fig. 1. First, the caption transition detection module filters out video frames without captions and detects transitions between caption texts. As a result, the video sequence is segmented into distinct image groups either sharing the same caption text or without captions. The image groups with captions are then fed into the caption-level enhancement module and the character-level enhancement module, resulting in an enhanced caption image. The caption-level enhancement module processes the entire caption area by the proposed mask operations, and the character-level enhancement module segments each caption into individual characters and removes noisy background residues. Finally, the OCR software accepts the caption image and generates text outputs. Our system only processes the caption region since its location is known as a priori for a particular TV program.

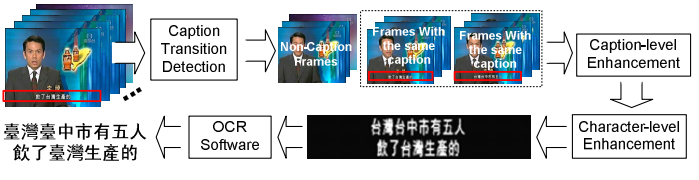


Fig. 1. The block diagram of our system

3 Caption Transition Detection

The most natural thought for caption transition detection is to compare a pair of immediate frames to examine whether or not they share the same caption content [4]. However, for a video retrieval task that involves processing of large numbers of videos, this frame-by-frame investigation is apparently not efficient. Thus we use a bi-search based caption transition detection approach that can significantly save computing time.

Since the same caption content usually lasts for a bunch of frames for stable viewing, we set up a searching step M and examine the difference between the current caption frame and the $(current + M)^{th}$ frame along a video sequence of N frames. Loop will continue if there is no major difference between these two frames (i.e. the difference does not exceed an empirical threshold). Otherwise (i.e. salient difference is found), bi-search will reduce the step by half each time and perform opposite direction search until reaches the first changing point (i.e. caption transition). Spatial difference [4] is calculated on a binary image pair after using morphological operations removing as many as possible background pixels. This approach effectively reduces the time complexity from $\mathcal{O}(N)$ (frame-by-frame difference) to $\mathcal{O}(\log N)$.

After caption transition detection, sequential image frames are automatically categorized into groups either sharing the same caption text or without captions. A representative image is then generated for each caption group for use in the enhancement step. We use minimum pixel search [4] to generate a representative image J_k for a caption group $\mathcal{G}_k$ that contains caption image I_t ($I_t \in \mathcal{G}_k$):

$$J_k = \min_{I_t \in \mathcal{G}_k} I_t. \tag{1}$$

4 Caption-Level Enhancement

4.1 Characteristics of Captions

Speech-synchronized captions always hold fixed positions and occupying a relatively stable area. Caption characters are often homochromous, surrounded by a thin edge with a different color, which acts as a 'fence' delineating the contour of individual characters. Although the edges can help humans easily distinguish captions from their background, computer may find this task hard since (1) character edges may vary a lot in color (eroded by background); (2) character fences are inevitably not always closed; (3) characters may mingle together. Cases (1) and (2) may induce a failure for edge detection based caption enhancement methods, while case (3) may lead to errors in character segmentation that is an important prerequisite in OCR. Some examples are shown in Fig. 2(a). These difficulties are caused by low resolution and lossy video compression. Morphological operations can be used to alleviate these problems with limited success [4].

For a sequence of frames with the same caption text, we found that background pixels have some nice features that can distinguish them from caption

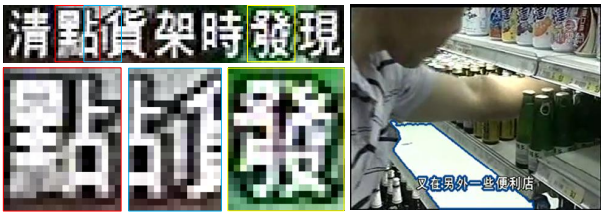


Fig. 2. Examples show the characteristics of captions

pixels. Background pixels are usually chromatic and change their intensity values among neighboring frames (background objects are in motion). Background areas with caption-like color are sometimes far bigger than a caption character, such as the big light area in Fig. 2 (b). Therefore, we propose to introduce three mask operations in caption enhancement based on these heuristics.

4.2 Mask Operations

We implement a *color mask*, a *changing mask* and a *region mask* to the caption region to remove non-caption pixels (i.e. background pixels), described as follows.

Color Mask. Previous approaches in caption enhancement mainly work on gray images and important color information is thus abandoned [4,6]. Since background pixels are usually chromatic while caption pixels are (near) homochromous, we adopt a color mask to distinguish them from each other. We first convert caption regions from RGB space to HSV space. Because saturation indicates the strength of color (from gray to color), we then use Otsu thresholding [7] to separate color pixels and monochromatic pixels apart in the saturation (S) scale by a logical color mask:

$$Mask^c = binary(I^S, Otsu(I^S)) \tag{2}$$

where I^S denotes the saturation space of caption HSV image and $Otsu(I^S)$ means using Otsu algorithm for thresholding on I^S. In the Otsu algorithm, the gray level histogram is divided in two classes and the optimal binarization threshold is selected that maximizes the between-class variance. Fig. 3 (a) shows an example of the original and its color-mask processed caption images. We can clearly see that the chromatic pixels in the background are converted to black after the color mask operation.

Changing Mask. Captions with the same text always keep still in a fixed position among frames and background objects usually move quite a lot. Bearing this in mind, we propose to use differences between immediate frames to remove the background pixels. We convert caption regions from RGB space to YIQ space, and apply the following XOR operation to the Y (luminance) scale between immediate frames.

$$Mask_t^m = xor(B_t, B_{t+1}), \tag{3}$$

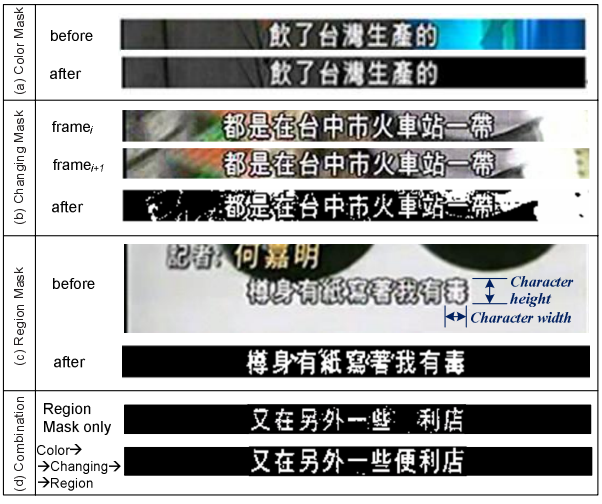


Fig. 3. Examples of (a) color, (b) changing, (c) region masks and (d) mask combination

$$B_t = binary(I_t^Y, Otsu(I_t^Y)). \tag{4}$$

I_t^Y denotes the Y space of tth caption YIQ image. Fig. 4(b) illustrates an example of enhanced caption region by the changing mask. We can clearly see that most of the background pixels are removed with intact caption pixels.

Region Mask. We notice that caption characters always hold a small area while background connected regions are always much bigger than that of a normal caption character. We thus calculate the areas of connected regions and pick out all those abnormally large regions. We use connected components labeling on the binary caption image, and remove the regions whose area is larger than the maximum number of pixels per character (a preset constant). Fig. 4 describes the generation of region mask. As shown in Fig. 3(c), region mask can effectively remove the big background regions that has a similar color with the captions.

Combination of Masks. As we pointed out earlier, the character edge is not always closed due to the low resolution and lossy compression. If we first apply the region mask to the binarized caption image, background pixels may 'break into' caption areas and thus remove these characters. To 'repair' the borders, we first use the color mask and changing mask operations to result in a binary image, on which region mask is calculated. This arrangement can greatly remove background pixels and disconnect the background and the caption in advance. As shown in Fig. 3(d), if we directly perform the region mask, character 便 will be removed since it is connected with the background. Instead, this will not happen if we use the binary result from the color mask and the changing mask to generate the region mask. Fig. 5 summarizes the steps of caption-level enhancement, where the enhancement is performed on the representative image J_k for each caption group $\mathcal{G}_k$.

```
RegionMaskGeneration(B)
Input  : binary caption image B with height of 3 × character_height
Output : region mask mask^r
begin
    LB = connected_components_labeling(B);
    STATS = sort(LB,'area');  % sort descending in area
    SB = LB(character_height, :); % shrink to normal height
    max_area = character_height × character_width
    mask^r = ones(size(SB)); % initialize region mask to 1
    for (i = 1 : length(STATS)) do
        if (STATS(i).area ≥ max_area) then
            Mask^r(find(SB == i)) = 0 % remove large area
        end
    end
end
```

Fig. 4. Pseudo-code for region mask generation

```
CaptionLevelEnhancement(J_k)
Input  : representative image J_k for caption group G_k
Output : enhanced caption B_k^Enh for G_k
begin
    J_k^Enh = J_k;  % initialization
    for (t = 1 : Frames(G_k)) do
        I_t^Enh = ImageFilt(I_t, Mask_t^c & Mask_t^m);
        B_t = binary(I_t^Enh, Otsu(I_t^Enh));
        Mask_t^r = RegionMaskGeneration(B_t);
        Mask_t = Mask_t^c & Mask_t^m & Mask_t^r % mask combine;
        J_k^Enh = ImageFilt(J_k^Enh, Mask_t); % enhancement
    end
    B_k^Enh = binary(J_k^Enh, Otsu(J_k^Enh)); % enhanced image
end
```

Fig. 5. Pseudo-code for caption level enhancement

5 Character-Level Enhancement

Character-level enhancement involves character segmentation and elaborate enhancement on individual characters. This module helps OCR software accurately locate individual characters, which is a preprocessing step in OCR before actually performing character recognition. We can tell from previous examples that some caption characters are mingled together, neither of which, if not separated, can be recognized. Moreover, although large area non-caption residues have been eliminated in caption-level enhancement, there are still some small residues remain in the caption region that will also result in a failure in OCR. For example, residues are around character 在 in Fig. 6. The character-level enhancement is

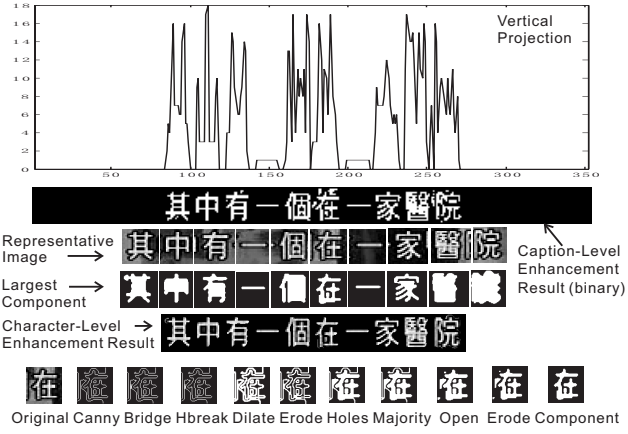


Fig. 6. Character segmentation and enhancement

thus adopted to segment individual characters and further remove residues to increase recognition accuracy.

We first adopt a vertical projection on the caption-level enhancement result, shown in Fig. 6. Then a simple thresholding and several heuristic rules are used to locate individual characters [4]. The segmentation process may induce inaccurate character borders. To ensure the integrity of each character, we expand the vertical span of each segmented character to extra two pixels at both sides. Although this may let the neighboring character strokes in, the final connected component analysis will remove these noises. We use a set of edge detection and morphological operations (e.g. canny, bridge, hbreak, fill holds, majority, shown in Fig. 6) to further process the individual characters. After the morphological operations, we label all the connected components in the binary character image to remove components that are too small as compared to a normal character. Finally, we use the binary image as a mask to remove residues and retain characters in the representative image, resulting in the character-level output.

6 Experiments

We have evaluated the proposed approach using Hong Kong TVB news programs that contain speech-synchronized captions of traditional Chinese characters. Videos are encoded in MPEG-2 with a resolution of 288×352. Six 30-min videos are adopted in the experiments, where one program is used for threshold empirical selections and the other five are used for testing.

In the test set of five videos, there are 3591 caption transitions. Our caption transition detection approach has successfully detected 3514 transitions, which leads to an accuracy rate of 97.86%. This performance is comparable to Tang's caption transition detection approach [4]. Our approach takes 14 mins in average to detect caption transitions in a 30-min video, while the frame-by-frame investigation needs 25 mins, tested on a Intel Core2 2.8GHz PC with 1G memory.

Table 1. OCR results on Captions

Method	Accuracy (%)
Representative image	10.50
Fixed binary threshold on representative image	16.11
Tang's caption enhancement approach [4]	81.63
Caption-level enhancement (proposed)	87.07
Caption and character enhancement (proposed)	89.24

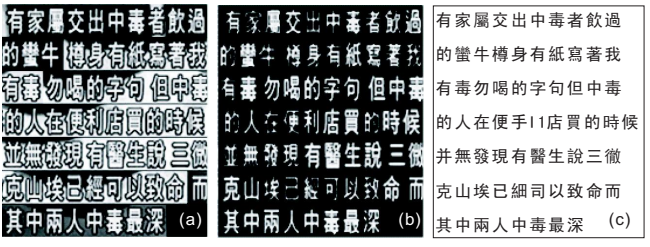


Fig. 7. (a) Original caption characters. (b) Caption characters enhanced by the proposed approach. (c) OCR result.

We use the TH-OCR software for character recognition experiments. As can be seen from Table 1, for the 23026 characters in the test set, we can only get character accuracy of 10.50% and 16.11% using the representative caption images and the simple thresholding, respectively. The proposed caption-level enhancement method achieves a high accuracy of 87.07% that significantly outperforms Tang's approach (81.63%). With the help of the character-level enhancement, the accuracy rate reaches 89.24%. As an example, Fig. 7 shows the OCR result of several captions. These caption characters share a similar or nearly the same color with their backgrounds. We can see that the troublesome background has been removed after our enhancement approach (Fig. 7 (b)), leading to a high OCR accuracy (Fig. 7 (c)).

7 Summary

This study attempts to improve the OCR accuracy of low resolution caption characters embedded in complex image backgrounds. Since the same caption text usually appears for multiple frames, we use a bi-search algorithm to improve the efficiency of caption transition detection. We make use of discriminative characteristics between captions and backgrounds and propose a combination of three mask operations to effectively enhance caption regions. We further segment individual characters and remove small background residues. Experiments show that our approach brings a high character OCR accuracy of 89.24%.

Acknowledgements

We thank Professor Helen Meng for constructive suggestions. We thank TVB (Hong Kong) for permitting us to use their news video data in the experiments. This work is partially supported by the Research Fund for the Doctoral Program of Higher Education in China (20070699015), the Natural Science Basic Research Plan in Shaanxi Province of China (2007F15), the NPU Aoxiang Star Plan (07XE0150) and the NPU Foundation for Fundamental Research (W018103).

References

1. TRECVID, http://www-nlpir.nist.gov/projects/t01v/
2. Xie, L., Liu, C., Meng, H.: Combined Use of Speaker- and Tone-Normalized Pitch Reset with Pause Duration for Automatic Story Segmentation in Mandarin Broadcast News. In: Proc. HLT-NAACL, pp. 193–196 (2007)
3. Xie, L., Zeng, J., Feng, W.: Multi-scale TextTiling for Automatic Story Segmentation in Chinese Broadcast News. In: Li, H., Liu, T., Ma, W.-Y., Sakai, T., Wong, K.-F., Zhou, G. (eds.) AIRS 2008. LNCS, vol. 4993, pp. 345–355. Springer, Heidelberg (2008)
4. Tang, X.-O., Gao, X.-B., Liu, J.-Z., Zhang, H.-J.: A Spatial-Temporal Approach for Video Caption Detection and Recognition. IEEE Trans. Neural Networks. 13(4), 961–971 (2002)
5. Lyu, M., Song, J., Cai, M.: A Comprehensive Method for Multilingual Video Text Detection, Locolization and Extraction. IEEE Trans. Circuits and Systems for Video Technology. 15, 243–255 (2005)
6. Sato, T., Kanade, T., Hughes, E.K., Smith, M.A., Satoh, S.: Video OCR: indexing digital news libraries by recognition of superimposed captions. Multimedia Systems 7, 385–395 (1999)
7. Otsu, N.: A threshold selection method from gray level histograms. IEEE Trans. Systems, Man and Cybernetics 9, 62–66 (1979)

A New Ear Recognition Approach for Personal Identification

Sepehr Attarchi[1], Masoud S. Nosrati[1], and Karim Faez[2]

[1] Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran
{threepehr,masoudnosrati}@aut.ac.ir
[2] Professor of Electrical Engineering Department, Amirkabir University of Technology, Tehran Iran
kfaez@aut.ac.ir

Abstract. Personal identification based on the ear structure is a new biometrics. In this paper, a new ear recognition method is proposed. For the purpose of segmentation, we apply the Canny edge detector to the ear image. Then the longest path in the edge image is extracted and selected as the outer boundary of the ear. By selecting the top, bottom, and left point of the detected boundary, we form a triangle with the selected points as its vertices. Further we calculate the barycenter of the triangle and select it as a reference point in all images. Then the ear region is extracted from the entire image using a predefined window centered at the reference point. For the recognition approach, we improve a previous method based on PCA by using 2D wavelet in three different directions and extracting three different coefficient matrices. Experimental results show the effectiveness of our proposed method.

Keywords: Ear segmentation, Canny edge detector, longest path, barycenter of the triangle, 2D Wavelet, PCA.

1 Introduction

Biometric personal identification is one of the most interesting topics in recent years. By increasing demands for the security in the societies, the old methods of personal identification are not reliable anymore. Ear recognition is a new class of biometrics and it has certain advantages over the more established biometrics; it has a rich and stable structure that is preserved from birth well into the old ages. The ear does not suffer from changes in facial expression, and is firmly fixed in the middle of the side of the head so that the immediate background is predictable, whereas face recognition usually requires the face to be captured against a controlled background. Collection does not have an associated hygiene issue, as may be the case with direct contact fingerprint scanning, and is not likely to cause anxiety, as may happen with iris and retina measurements. The ear is large compared with the iris, retina, and fingerprint and therefore is more easily captured [1].

In recent years, some methods have been developed for ear recognition. Iannarelli used a manual method in his research. He examined over 10000 ears and proved their uniqueness, based on measuring the distances between specific points of the ears [2]

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 356–363, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The main problem in using this method is detecting those specific key points automatically.

Burge and Burger [3, 4] proposed a method based on neighborhood graph and Voronoi diagrams of the detected edges. They did not give any results from their experiments. Their method is not enough reliable because of their feature extraction based on ear edges which can be affected by noise. Hurley et al. [1] applied force field transform to ear images in order to find energy lines, wells and channels.

Choras [5] proposed a method based on geometrical feature extraction. After edge detection, he creates a set of circles with the center in the centroid. These circles are crossed by the ear edges and he counts the number of intersecting pixels. His feature vector consists of all the radii with the corresponding number of pixels belonging to each radius and the sum of all the distances between those pixels.

In this paper, we propose a new approach for ear recognition. For the purpose of segmentation, Canny edge detector is used to obtain the edges of the ear image. Then the longest path in the edge image is extracted and selected as the outer boundary of the ear. By selecting the top, bottom, and left point of the detected boundary, we form a triangle with the selected points as its vertices. Further we calculate the barycenter of the triangle and select it as a reference point in all images. Then the ear region is extracted using a predefined window centered at the reference point. To calculate the predefined window's dimensions, the mean ear template is computed by using the top, bottom and the left points of the extracted ear boundary. For the recognition approach, we improve a previous method based on PCA by using 2D wavelet. After intensity and orientation normalization of the ear image, we apply 2D wavelet to the normalized image in three different directions and extract three different coefficient matrices (Horizontal, Vertical and Diagonal). After this, we integrate these matrices and extract a new feature matrix.

The rest of the paper is organized as follows: our segmentation approach is introduced in Section 2. The normalization procedure is presented in Section 3. Our proposed feature extraction is given in section 4. The experimental results are shown Section 5; and finally the conclusions are given in Section 6.

2 Ear Segmentation

Template Matching is a very common way for ear segmentation [1]. For this purpose, the mean image of ear is computed. The computed mean image is slid across the original image and the correlation between the image and the template is computed in each step. Using a template for ear segmentation is not a suitable approach for ear segmentation because we are not able to locate the ear if the original image has been rotated or if the original image has a different size in compare of our template. For the purpose of segmentation, we use the Canny operator. 8 are selected as the standard deviation of the Gaussian filter in the Canny algorithm. By assessing the edge image, our experiments show that the longest path in the edge image is the outer boundary of the ear. Therefore we select the longest path of the edge image as the border of the ear. Fig. 1 shows some instances for the described method.

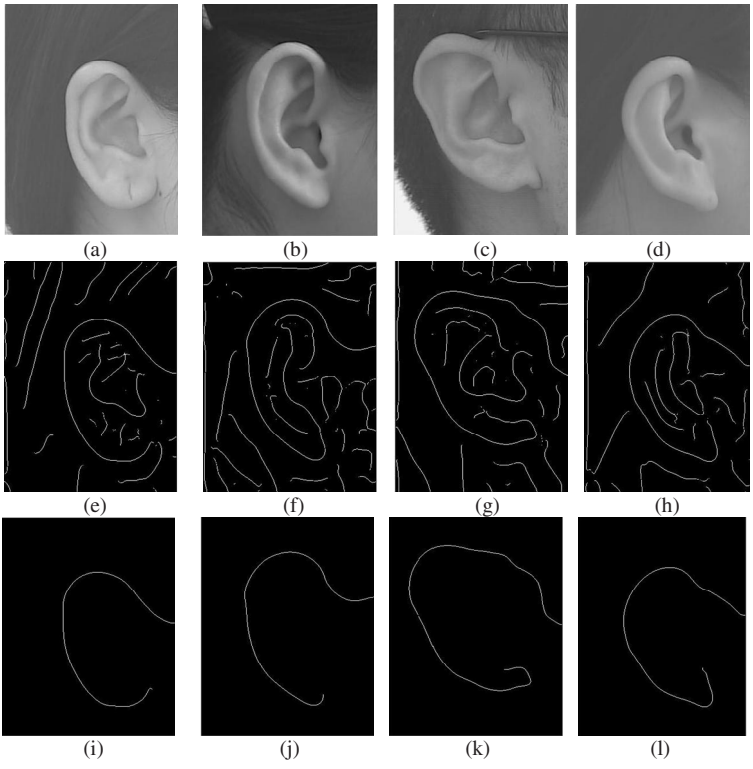


Fig. 1. Original ear images (a, b, c, d). Ear images after applying the Canny edge detector (e, f, g, h). The Longest path as the boundary of each ear image (i, j, k, l).

After finding the longest path, the top, bottom and the left points of the path are detected. Then we form a triangle using these points as the vertices triangle. By computing the barycenter of triangle, we choose this point as a reference point in all images. We extract the ear region with a predefined fixed dimension window centered at the reference point. To calculate the predefined window, the mean ear template is computed from all images in dataset by using the top, bottom and the left points of the extracted ear boundary as fixed points of the template borders. Fig. 2 shows the results of our proposed segmentation method.

3 Normalization

In order to reduce the destructive effects of illumination, we normalize a given segmented ear image *I*. Let *I(i,j)* denotes the intensity of the image at pixel *(i,j)*. If *M* and *Var* indicate the mean and variance of the input image respectively, the intensity of pixel *(i,j)* in the normalized image is obtained from the following formula [6]:

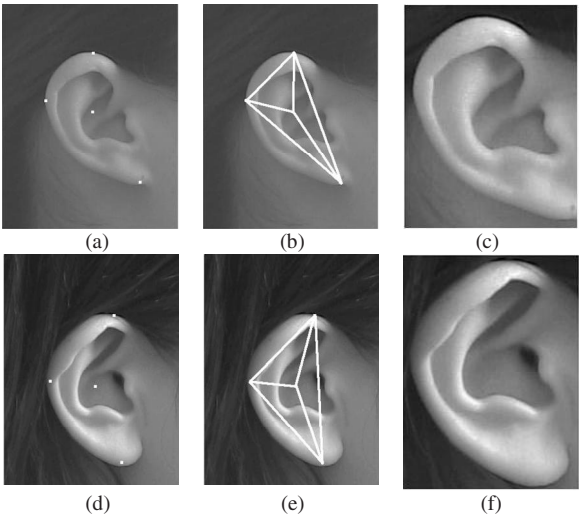


Fig. 2. Top, bottom, and the left points of the selected outer boundary with the computed barycenters (a, b, d, e). Extracted region of ears using a predefined window(c, f).

$$N(i,j)=\begin{cases} M_0+\sqrt{\dfrac{Var_0(I(i,j)-M)^2}{Var}} & I(i,j)>M \\ M_0-\sqrt{\dfrac{Var_0(I(i,j)-M)^2}{Var}} & otherwise \end{cases} \tag{1}$$

where M_0 and Var_0 are desired mean and variance for the normalized image. Fig. 3 shows the result of intensity normalization. To make our proposed method robust to the rotation, we perform an orientation normalization procedure. After finding the longest path, we detect the line from the top of the ear to the bottom. This line is used to find the orientation in x-y plane. For this purpose, we calculate the rotation angle using equation (2).

$$\theta=\tan^{-1}\frac{y}{x} \tag{2}$$

where x and y are the vertical and horizontal distance between the top and the bottom points respectively. We rotate the image by multiplying the rotation matrix by the image as shown in equation 3. Fig. 4 shows the steps of normalization.

$$\begin{bmatrix} x_{new} \\ y_{new} \end{bmatrix}=\begin{bmatrix} Cos(90-\theta) & Sin(90-\theta) \\ -Sin(90-\theta) & Cos(90-\theta) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \tag{3}$$

(a) (b) (c) (d)

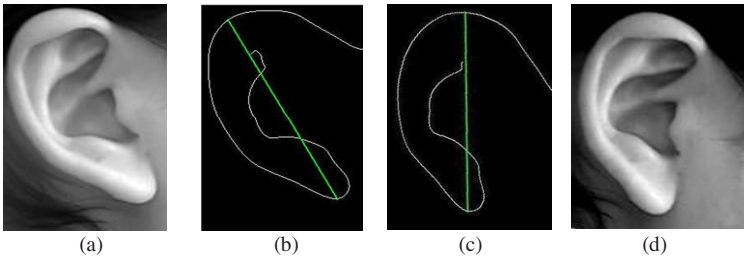
Fig. 3. Extracted region of ear be fore intensity normalization (a, c). Extracted region of ear After intensity normalization (b, d).

4 Feature Extraction Using 2D Wavelet

The most suitable information for feature extraction locates on the edges and high frequency regions of the ear. For extracting these features we need to consider the ear image in frequency domain. 2D wavelet let us to have the frequency and spatial domain information simultaneously. In fact, the strength of high frequency components can be interpreted as a measure of changes at the spatial domain. In addition by applying 2D wavelet, we can analyze the image in different directions. These motivate us to analyze the ear image in scale-frequency domain to extract the important features. With each wavelet type of this class, there is a scaling function (also called father wavelet) which generates an orthogonal multi resolution analysis. We apply 2D-wavelet in three directions (Horizontal, vertical and diagonal) separately to find three independent features. For evaluating the effects of each wavelet coefficients (Horizontal, Vertical and Diagonal matrices) on the recognition rate, we use the PCA[1] method for each of these coefficients separately.

The results show that horizontal coefficients matrix has larger recognition rate than vertical coefficients and diagonal coefficients matrices. Table 1 shows the recognition rate after using the PCA method for each of these matrices on two ear databases, USTB database2 [7] and Carreira-Perpinan database [8].

After this, we integrate the matrices of wavelet coefficients in the three directions (Horizontal, vertical and Diagonal) by adding the absolute value of these matrices:



(a) (b) (c) (d)

Fig. 4. Extracted ear region (with the edge image) before orientation normalization (a, b). Extracted ear region (with the edge image) after orientation normalization (c, d).

[1] Principle Component Analysis.

where α, β and γ are constants. This combination allows us to consider the changes in the ear images from three basic directions.

So we choose α, β and γ related to this recognition rate. Since the recognition rate of horizontal wavelet components are higher than the others, its weight should be greater. Here we choose $\alpha = .36$, $\beta = 0.33$ and $\gamma = 0.31$. Experimental results show the effectiveness of our proposed method for the recognition purpose.

$$F = \alpha \ |LH| + \beta \ |HL| + \gamma \ |HH| \tag{4}$$

5 Experimental Results

We tested the proposed method in both segmentation and recognition approaches using two ear databases with different noise level, contrast and average brightness. We used USTB database2 (which includes 308 ear images from 77 persons) and Carreira-Perpinan database (which includes 102 ear images from 17 persons). First we describe our segmentation results and then we come to our recognition experiments. As mentioned before, template matching is a common approach for ear segmentation in many studies. Also, to the best of our knowledge, there are a few different approaches for ear segmentation. Also there are some manual segmentation methods for ear recognition [9]. Therefore in order to evaluate the accuracy of our proposed segmentation method, we designed and implemented a template matching method. To choose a template for the template matching procedure, we selected the same dimensions for the template as we used for fixed dimension window in our proposed method. Then the mean image of ear is computed using all images in database. The window was slid across the original image and the correlation between the image and the window was computed in each step. Table 2 shows the segmentation results of the proposed method in comparison with implemented template matching procedure. All the calculation was implemented using Matlab 7 on an Intel Pentium IV 2.66G processor with 512 MB memory. The execution times for ear segmentation for both databases are shown in table 3. As it is clear our proposed method shows better performance in both accuracy and execution time of the segmentation. In the recognition procedure, for training the PCA in the USTB database, we choose two images (from each person) for training and two images for the test. In the second database, for the first experiment, we choose 3 images (from each person) for training and 3 images for test and for second one, we choose 4 images (from each person) for training and 2 images for test. Table 4 shows the recognition results of our proposed method [10]. As it is clear our recognition method outperforms the previous PCA based methods.

Table 1. Recognition rate, using PCA for three wavelet features in three directions

Wavelet feature	Horizontal	Vertical	Diagonal
Average recognition rate (%)	90.19	86.27	82.35

Table 2. Comparing the segmentation results of our proposed method with Template Matching

	Template Matching		Accuracy (Template Matching)	**Proposed Method**		Accuracy (Proposed Method)
	Number of correct segmented images	Number of incorrect segmented images		Number of correct segmented images	Number of incorrect segmented images	
USTB DB	284	24	92.2%	302	6	98.05%
Carreira-Perpinan DB	94	8	92.15%	99	3	97.05%

Table 3. Time comparison between the proposed method and Template Matching

	Time (s) Template Matching	Time (s) Proposed Method
USTB DB	18.82	0.97
Carreira-Perpinan DB	0.42	0.031

Table 4. Comparing the previously introduced PCA method and proposed method

	Previously introduced PCA method [9]	**Our proposed method** (Training PCA by wavelet coefficients)
USTB DB (2 images for training)	84.3%	90.5%
Carreira-Perpinan DB (3 images for training)	86.27%	95.05%
Carreira-Perpinan DB (4 images for training)	88.22%	97.05%

6 Conclusions

In this paper, we proposed a new approach for ear recognition. As the segmentation approach, we used the Canny edge detector to obtain the edges of the ear image. Then we selected the longest path of the edge image as the outer boundary of the ear. By selecting the top, bottom, and left point of the detected boundary, we formed a triangle with the selected points as its vertices. We calculated the barycenter of the triangle and used it as a reference point in all images. Then we extracted the ear region with a predefined window centered at the reference point. For the recognition approach, we improved a previous method based on PCA by using 2D wavelet in three different directions. We extracted three different coefficient matrices. By integrating these matrices with different weights, we formed a new feature matrix. We achieved the correct segmentation rate of 98.05% and 97.05% for two different databases. Also we reached the recognition rate of 90.5% and 95.05% for two sets of database.

References

1. David, J., Hurley, M.S., Nixon, J.N.C.: Force Field Feature Extraction for Ear Biometrics. Computer Vision and Image Understanding 98, 491–512 (2005)
2. Iannarelli, A.: Ear Identification: Forensic Identification Series. Paramont Publishing Company, California (1989)
3. Burge, M., Burger, W.: Ear Biometrics. Johannes Kepler University, Linz, Austria (1999)
4. Burge, M., Burger, W.: Ear Biometrics for Machine Vision. In: 21 Workshop of the Austrian Association for Pattern Recognition, Hallstatt (1997)
5. Michał, C.: Ear Biometrics Based on Geometrical Feature Extraction. Electronic Letters on Computer Vision and Image Analysis 5(3), 84–95 (2005)
6. Wang, S., Wang, Y.: Fingerprint Enhancement in the Singular Point Area. IEEE signal Processing Letters 11(1) (2004)
7. Yuan, L., Mu, Z., Xu, Z.: Using Ear Biometrics for Personal Recognition. In: Li, S.Z., Sun, Z., Tan, T., Pankanti, S., Chollet, G., Zhang, D. (eds.) IWBRS 2005. LNCS, vol. 3781, pp. 221–228. Springer, Heidelberg (2005)
8. Carreira-Perpinan.: Compression Neural Networks for Feature Extraction: Application to Human Recognition From Ear Images. M.Sc thesis, Faculty of Informatics, Technical University of Madrid, Spain (1995) (in Spanish)
9. Kevin, P., Bowyer, W.: Ear Biometrics Using 2D and 3D images. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005) 1063-6919/05 (2005)
10. Masoud, S., Nosrati, K.F., Farhad, F.: Using 2D Wavelet and Principal Component Analysis for Personal Identification Based on 2D Ear Structure. In: International Conference on Intelligent and Advanced Systems, ICIAS 2007, Kuala Lumpur, Malaysia (November 2007)

An FSM-Based Approach for Malicious Code Detection Using the Self-Relocation Gene

Yu Zhang[1], Tao Li[1], Jia Sun[2], and Renchao Qin[1]

[1] College of Computer Science, Sichuan University, Chengdu 610065, China
bullzhangyu@yahoo.com.cn
[2] Department of Humanism Educations, Huaihua University, Huaihua 418000, China

Abstract. Most information attacks on the Internet are perpetrated by deploying malicious codes, which results in destructive epidemics. The correct execution of viruses and worms throughout the Internet is accomplished by self-relocation. Self-relocation is a built-in mechanism in most malicious codes, allowing them to get the base address of the host program to correctly infect it. Since most legitimate computer programs do not need self-relocate themselves, the detection of malicious codes could be reduced to the detection of the various mutations of the self-relocation gene. This study presents such a detection mechanism based on finite state machine theory for both known and previously unknown malicious executable codes. It does not rely on signature-screening of known viruses, but instead it detects self-relocation attempt of the suspicious executable code. The experiments were conducted and the results indicate that the proposed approach has better ability to detect known and previously unknown malicious executable codes than other methods.

Keywords: malicious code detection, self-relocation gene, finite-state machine.

1 Introduction

Malicious code has existed since the dawn of computing and it has become increasingly prevalent in recent years [1][2]. Malicious codes such as Trojan horses, worms, and viruses pose an increasing risk to integrity, confidentiality, availability of information, and potentially administrative control. They cause loss of valuable data and cost an enormous amount in wasted effort to restore damaged data [3].

To effectively protect information systems against malicious codes attack, the researchers have developed a variety of detection techniques. Those techniques used for detecting malicious codes can be categorized broadly into three categories: anomaly-based detection, specification-based detection and signature-based detection [4][5][6][7]. Anomaly-based detection techniques use the knowledge of what constitutes normal behavior to decide the maliciousness of a program under inspection. Specification-based techniques leverage some specification of what is valid behavior to decide the maliciousness of a program under inspection. Signature-based detection uses its characterization of what is known to be malicious to decide the maliciousness of a program under inspection [8][9][10][11]. The fundamental limitation of anomaly-based detection is its high false-positive rate and the time complexity in the training

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 364–371, 2008.
© Springer-Verlag Berlin Heidelberg 2008

phase. The main limitation of specification-based detection is that it is difficult to specify completely and accurately the entire set of valid behaviors. The primary drawback of signature-based detection is that it cannot detect zero-day attack, for which there is no corresponding signature stored in the database [12].

In order to improve the performance of currently malicious code detection, we proposed an FSM-based approach for malicious code detection using the self-relocation gene. We have studied various malicious codes and found that the malicious codes cannot correctly execute without self-relocation [13] and thus attack objectives cannot be fully achieved. Self-relocation, which is uncommon in legitimate programs, is vital to the execution of computer viruses allowing them to obtain the base address of the host codes. Since most legitimate computer programs do not need self-relocate, the detection of malicious codes could be reduced to the detection of the various mutations of the self-relocation gene. The experiments were conducted and the results show that this approach has better efficiency in the detection of known and previously unknown malicious codes than the others.

In the following sections, we first describe the definition of the self-relocation gene in section 2. Then we introduce the mechanism of the self-relocation gene detection in section 3. Section 4 shows the implementation and experiment results. We state our conclusion and future work in Section 5.

2 Definition of the Self-Relocation Gene

Self-relocation, which is uncommon in legitimate programs, is important to the execution of malicious codes allowing them to obtain the base address of the host codes thus creating computer epidemics and maximizing the effectiveness of the attack. The self-relocation gene (SRG) can be regarded as a specific sequence of commands that cause the malicious code to relocate itself and thus to execute correctly in host programs. The typical self-relocation gene-- E8000000005B81EB5C1A4000-- in the malicious executable codes is shown in Figure 1.

Offset	Opcode	Instruction
.text:00401A57	E800000000	call $+5
.text:00401A5C	loc_401A5C:	
.text:00401A5C	5B	pop ebx
.text:00401A5D	81 EB 5C 1A 40 00	sub ebx, offset loc_401A5C

Fig. 1. The typical self-relocation gene

Self-relocation can be accomplished in various ways relying on malicious code writers' techniques. Since the most sophisticated and versatile viruses are still implemented in assembly language and assembled into executable files, the self-relocation gene could only be possibly obtained in search of low-level machine codes.

The SRG is usually contained within the code sequence of malicious code and it could also be dispersed throughout that sequence to become a variety of SRG mutations in order to anti-disassemble. Since none of the machine command alone can be considered malicious, only the particular sequences of commands can construct the

SRG. Therefore, the SRG is described using the concept of building blocks, where each block performs a part of the self-relocation. This concept is illustrated in Figure 2. The upper blocks of the self-relocation gene base on the lower blocks. Most of the building blocks involved in malicious self-relocation activity can individually be included in any program for legitimate reasons. Only when integrated into larger structures based on their inter-functional relationships, these building blocks can be regarded as attempts to self-relocate.

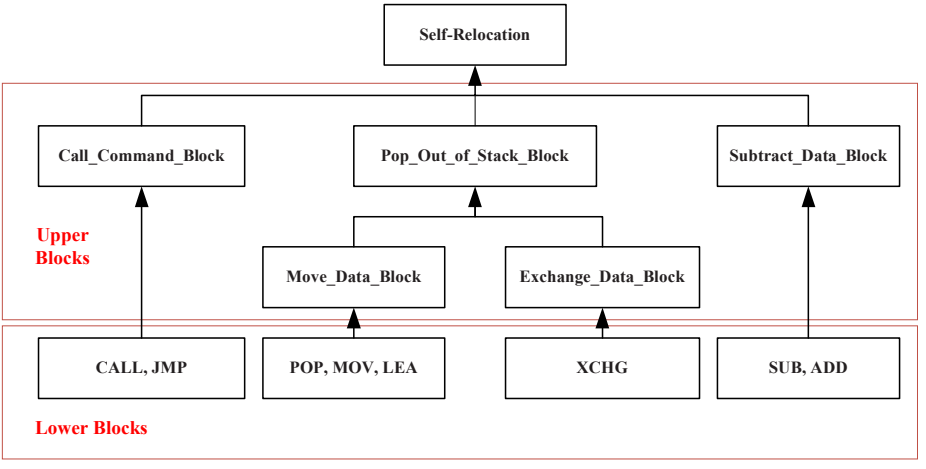


Fig. 2. The SRG structure

The SRG consists of such blocks in various ways. Therefore, its structure can be viewed as a regular sentence being built up by concatenating phrases, where phrases are built up by concatenating words, and words are built up by concatenating characters.

The aim of applying such a syntactic approach to describe the SRG is to facilitate the recognition of sub-patterns. This suggests the recognition of smaller building blocks first, establishing their relevance and contribution to the self-relocation, and then considering the next sub-pattern. The syntactic description of the SRG provides a capability for describing and detecting large sets of complex patterns by using small subsets of simple pattern primitives. It is also possible to apply such a description any number of times to express the basic structures of lots of gene mutations.

Following the concept of syntactic description, the SRG structure could be represented using the grammar definition notations [14]:

$$G = \{V_N, V_T, P, S\} \tag{1}$$

where

G is the self-relocation gene; V_N is the non-terminal variable; V_T is the terminal variable; P is the finite set of rules; S is the starting point of the gene.

The non-terminal variable V_N in the expression above can be defined as:

$$V_N = \left\{ \begin{array}{l} \text{<Self_Relocation_Gene>, <Call_Command_Block>,} \\ \text{<Pop_Out_of_Stack_Block>, <Move_Data_Block>,} \\ \text{<Exchange_Data_Block>, <Subtract_Data_Block>} \end{array} \right\} \quad (2)$$

The terminal variable V_T represents the SRG sequence:

$$V_T = \{\text{CALL, JMP, POP, MOV, LEA, XCHG, SUB, ADD}\} \quad (3)$$

The complete vocabulary *V(G)* is composed of V_N and V_T , that is, $V(G)= V_N \cup V_T$ and $V_N \cap V_T = \varnothing$.

P, the set of rules, is expressed as $\alpha \rightarrow \beta$, where α and β are any blocks in V_N, and α and β interconnects in *V(G)*.

$S \in V_N$ represents the starting point in V_N, which is equal to the < Self_Relocation_Gene >.

3 The SRG Detection Mechanism

Since the SRG structure is defined in terms of sub-patterns similar to the structure of a sentence with its phrases, words and characters, the finite-state machine theory for text recognition is applicable for the SRG detection. Finite-state machine is a model of computation consisting of a set of state, a start state, an input alphabet, and a transition function that maps input symbols and current states to a next state. Computation begins in the start state with an input string. It changes to new states depending on the transition function.

Mathematically, a finite-state machine M is a quintuple,

$$M= (Q, \Sigma, \delta, q_0, F) \quad (4)$$

where

Q is a finite, non-empty set of states; Σ is the input alphabet (a finite, non-empty set of symbols); δ is the state-transition function: $\delta: Q \times \Sigma \rightarrow Q$; $q_0 \in Q$ is an initial state, an element of *Q*; $F \subseteq Q$ is the set of final states.

We can define a finite-state machine $M=\{V_T, V_N \cup \{T\}, \delta, S, F\}$ with *T(M) = L(G)*, if $G =\{V_N, V_T, P, S\}$ defined above, is a finite-state SRG expression. Since P always contains relation rule for S when detecting a SRG, the set of final states *F* contains *S*, that is, $F =\{S, T\}$. Therefore, the finite-state machine can be constructed for SRG detection purposes, all relocation combination that are accepted by the finite-state machine are in the state space of a phrase-structure language defined as L(G). This language is to be created by the SRG grammar in the following way:

$$L(G) = \{x \mid x \in V_T\} \quad (5)$$

$$S \xrightarrow[G]{} x \quad (6)$$

where

x is a relocation building block; $S \xrightarrow[G]{} x$ implies that *x* is directly derived by another building block *S*, and that both *x* and *S* follow the rule *P* by yielding

$S = \omega_1 \alpha \omega_2$ and $x = \omega_1 \beta \omega_2$, where $\alpha \rightarrow \beta$ is defined in P, and ω_1 and ω_2 are the elements in V_T.

During the SRG detection process, every machine code comes into the self-relocation detector, where it goes through two detection mechanisms. Just like the SRG is formed from many different building blocks, the detection mechanism first observes every lower block separately, then if reaching the upper block of the SRG and all upper blocks obtaining, it will declare the alarm state and make decisions of malicious codes. The brief diagram of detection algorithm is shown in Figure 4.

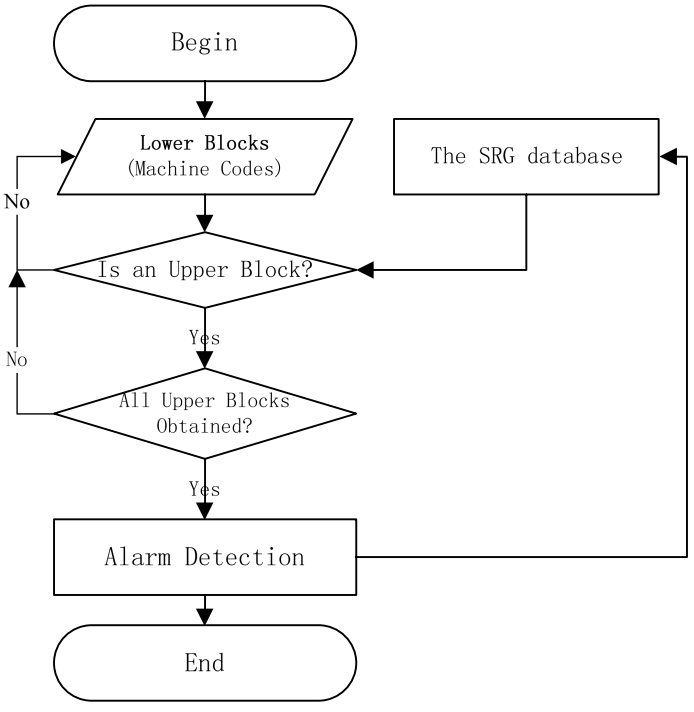


Fig. 4. The detection algorithm

When detecting the suspicious machine codes, the SRG database determines whether or not this code can match any other lower level blocks to form an upper level block structure. If a new upper block is finally formed, the detector will alarm the suspicious code. Then the new SRG of that suspicious machine code will be put into the SRG database only if it is not there. Therefore, the detector has the ability to self-learn when detecting the suspicious codes.

4 Experiment and Results

The experiment was conducted in the computer virus and anti-virus laboratory, computer network and information security Institute of Sichuan University. There is no

benchmark data set available for the malicious code detection unlike intrusion detection. Our data sets include 100 viruses collected from the website VX Heavens [15] and 500 benign executables taken from system32 folder in Windows. The main goal of the experiment is to test the detection rate of known and unknown viruses and false-positive rate of the normal files. The experimental results are shown in Figure 5. Figure 5 shows that the detection rate of the malicious codes is 97%, the omitting rate is 3%, and the false-positive rate (misidentification of legitimate programs as viruses) is only 3.6%. This exciting experimental result is due to the FSM-based detector that detects the lower level machine codes combing with the self-relocation gene database to form upper level blocks, which makes it effectively detect malicious codes with lower false-positive rate. The detectors can detect most of Windows PE viruses except the packed and encrypted ones. Since the packed and encrypted viruses were protected by the shell and the inner binary codes including the relocation part were disturbed with encryption, the detection of them will result in the omitting rate. Meanwhile, since there are no differences between data and codes in the computer with Intel architecture, some legitimate system programs that may include some data equivalent to the relocation code lead to false-positive rate.

Fig. 5. The experiment results of our approach

In order to test the performance of the proposed approach, we conducted the related comparison experiments with the currently most mature antivirus product Kaspersky7.0 [16] and with the presently most used method APIs [17], which uses API sequence to detect malicious codes.

The comparison experiments results are shown in Figure 6 that the detection rate of our proposed approach is 97%, Kaspersky 88%, and APIs 65%. Since most viruses do not involve APIs in IAT (Import Address Table), the detection of APIs can result in higher omitting rate. And since the Kaspersky antivirus technology is signature-based, it can only detect known computer viruses. As a result, it has the higher detection rate of known viruses but lower detection rate of previously unknown viruses. Unlike them, our approach is a heuristic technique, which can effectively detect known and previously unknown viruses. The results indicate that our proposed approach has higher detection rate than the others, and efficiently testify the validity of our proposed approach.

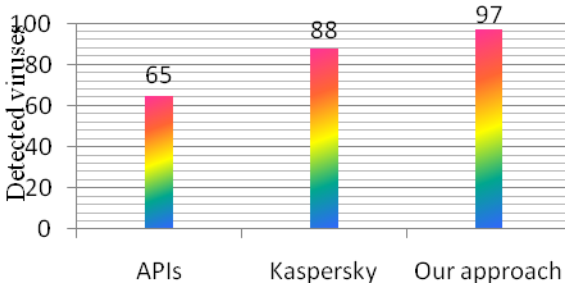


Fig. 6. The comparison experiments results

5 Conclusion and Future Work

In this study we proposed a novel FSM-based approach for malicious code detection using the self-relocation gene. The reason for choosing the mechanism of self-relocation as the detection criteria is that most legitimate codes do not need self-relocate while running, but malicious codes have to self-relocate themselves while executing. The main strength of the proposed approach is its ability to detect previously unknown viruses with a very low false-positive rate. Moreover, it is independent of the style of the programming language used because the detection is done at a very low level with machine codes. The experimental results show that the proposed approach not only has a high detection rate, low false-positive rate and low omitting rate, but also its efficiency is better than other methods.

The future work will involve extending this approach to detect packed and encrypted malicious codes, which are presently the most dangerous threat to the computer network systems.

Acknowledgement. This work was sponsored by the National Natural Science Foundation of China (60573130), National 863 project of China (2006AA01Z435), and the New Century Excellent Expert Program of Ministry of Education of China (NCET-04-0870).The authors acknowledge the support of computer virus and antivirus Lab of Sichuan University and thank the VX Heavens!

References

1. Ford, R., Spafford, E.H.: Happy Birthday, Dear Viruses. Science 317(5835), 210–211 (2007)
2. Trilling, S., Nachenberg, C.: The Future of Malware. In: EICAR Proceedings (1999)
3. Kumar, S., Spafford, E.H.: A Generic Virus Scanner in C++. In: Proceedings of the 8th Computer Security Applications Conference, pp. 210–219 (1992)
4. Idika, N., Mathur, A.: A Survey of Malware Detection Techniques (2007), `http://www.serc.net/report/tr286.pdf`
5. Skormin, V., Volynkin, A., Summerville, D., Moronski, J.: In the Search of the “Gene of Self-replication” in Malicious Codes. In: Proceedings of IEEE Workshop on Information Assurance and Security, pp. 193–200 (2005)

6. Skormin, V., Summerville, D., Moronski, J.: Detecting Malicious Codes by the Presence of Their Gene of Self-Replication. In: Gorodetsky, V., Popyack, L.J., Skormin, V.A. (eds.) MMM-ACNS 2003. LNCS, vol. 2776, pp. 195–205. Springer, Heidelberg (2003)
7. Summerville, D., Skormin, V., Volynkin, A., Moronski, J.: Prevention of Information Attacks by Run-Time Detection of Self-replication in Computer Codes. In: Gorodetsky, V., Kotenko, I., Skormin, V.A. (eds.) MMM-ACNS 2005. LNCS, vol. 3685, pp. 54–75. Springer, Heidelberg (2005)
8. Ellis, D.R., Aiken, J.G., Attwood, K.S., Tenaglia, S.D.: A Behavioral Approach to Worm Detection. In: Proceedings of the 2004 ACM Workshop on Rapid Malcode, pp. 43–53 (2004)
9. Moskovitch, P., Nissim, N., Elovici, Y.: Malicious Code Detection and Acquisition Using Active Learning. IEEE Intelligence and Security Informatics, 371 (2007)
10. Elovici, Y., Shabtai, A., Moskovitch, R., Tahan, G., Glezer, C.: Applying Machine Learning Techniques for Detection of Malicious Code in Network Traffic. In: Proceedings of the 30th Annual German Conference on Artificial Intelligence, pp. 44–50 (2007)
11. Zhang, Y., Li, T., Qin, R.: A Dynamic Immunity-based Model for Computer Virus Detection. In: Proceedings of International Symposium on Information Processing, pp. 515–519 (2008)
12. Kolter, J., Maloof, M.: Learning to Detect Malicious Executables in the wild. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 470–478 (2004)
13. Christodorescu, M., Jha, S.: Static Analysis of Executables to Detect Malicious Patterns. In: Proceedings of the 12th USENIX Security Symposium, pp. 169–186 (2003)
14. Michael, S.: Introduction to the Theory of Computation, 2nd edn. Thomson Course Technology, Boston (2006)
15. VX Heavens, `http://vx.netlux.org`
16. Kaspersky Lab, `http://www.kaspersky.com`
17. Xu, J., Sung, A., Chavez, P., Mukkamala, S.: Polymorphic Malicious Executable Scanner by API sequence analysis. In: Fourth International Conference on Hybrid Intelligent Systems, pp. 378–383 (2006)

WTSPMiner: Efficiently Mining Weighted Sequential Patterns from Directed Graph Traversals*

Runian Geng[1,2], Xiangjun Dong[2], He Jiang[2], and Wenbo Xu[1]

[1] School of Information Technology, Jiangnan University
[2] School of Information Science and Technology, Shandong Institute of Light Industry
gengrnn@163.com, d-xj@163.com, jianghe@sdili.edu.cn,
xwb_sytu@hotmail.com

Abstract. Data mining on graph traversals has been an active research during recent years. However, traditional traversal patterns mining algorithms only considered un-weighted traversals, and they are most un-sequential patterns mining algorithms. Based on the property that the traversal pattern and the items in it are consecutive, this paper regards a traversal patter as a particular sequential pattern and proposes a new algorithm, called *WTSPMiner* (*W*eighted *T*raversal-based *S*equential *P*atterns *Miner*), to mine weighted sequential patterns from traversals on weighted directed graph. Adopting an improved weighted prefix-projected pattern growth approach, *WTSPMiner* decompose the task of mining original sequence database into a series of smaller tasks of mining locally projected database so as to efficiently discover fewer but important weighted sequential patterns. Comprehensive experimental results show that *WTSPMiner* is efficient and scalable for finding weighted sequential patterns from weighted graph traversals.

Keywords: data mining, weighted directed graph, traversal pattern mining, weighted sequential pattern mining.

1 Introduction

Data mining on graph traversals has been an active research field during recent years due to its broad applications. One example of it is Web whose structure can be modeled as a *DG* (*D*irected *G*raph) in which the vertices represent Web pages, and the edges represent hyperlinks between the pages. Capturing user access patterns in such environments is referred to as mining traversal patterns [1]. Traditional traversal patterns mining algorithms only considered unweighted traversals [1,2]. To reflect importance of Web elements, we can assign a weight to each site. For example, the weights assigned to vertices can reflect the importance difference of each Web page's content or the cost time user stay on the page. Each edge can be assigned with a

* This work was supported in part by the Natural Science Fund of Shandong Province (No.Y2007G25), the Excellent Young Scientist Foundation of Shandong Province, China (No.2006BS01017), and the Scientific Research Development Project of Shandong Provincial Education Department, China (No. J06N06).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 372–379, 2008.
© Springer-Verlag Berlin Heidelberg 2008

weight standing for the cost time of traversal between two pages. Accordingly we can model the structure of Web with a *WDG* (*W*eighted *D*irected *G*raph). Sequential pattern mining has been an active research topic in data mining. Most sequential pattern mining algorithms [3,7] use an Apriori-based pruning strategy to reduce the number of candidates. However, they are ineffective since they will repeatedly scan the sequence database and generate a large number of candidates to be checked. To overcome this shortcomings, the algorithms of mining sequential pattern by growth methods [4,5] have been presented. However, they do not consider the weight constraint. In addition, most previous weighted mining algorithms [8,9] do not concern in mining frequent sequence from traversals.

In this paper, we extended previous works by attaching weights to the traversals and introduced a new effective and scalable algorithm called *WTSPMiner* (*W*eighted *T*raversal-based *S*equential *P*atterns *M*iner) to discover weighted sequential patterns from traversals on *WDG*. Adopting a 'divide-and-conquer' strategy, *WTSPMiner* pushes the weight constraint into the mining process so as to efficiently discover fewer but more important weighted traversal patterns.

The remainder of this paper is organized as follows. Section 2 gives the related definitions of problem. *WTSPMiner* algorithm is proposed in Section 3. In Section 4, we present the experimental results. Finally, Section 5 gives the conclusion.

2 Problem Statement

Let $I = \{i_1, i_2 \ldots i_n\}$ be a set of all items. An itemset is a subset of I. A sequence $S = \langle s_1, s_2 \ldots s_m \rangle$, where itemset $s_j \subseteq I$, $j \in \{1 \ldots m\}$. s_j is also called an element of S, and denoted as $(i_{j_1} i_{j_2} \ldots i_{j_l})$, where i_{j_k} is an item in s_i, l is called size of s_j. The brackets are omitted if an element has only one item. An item can occur at most once in an element of a sequence, but can occur multiple times in different elements of a sequence. The size of S, $|S|$, is the number of itemsets or elements in S. The size of s_j, $|s_j|$, is the number of items in s_j. The length of sequence S, $l(S)$, is the sum of all elements size in it, i.e., $l(S) = \sum_{j=1}^{m} |s_j|$.

2.1 Related Definitions of Weighted Sequential Pattern Mining

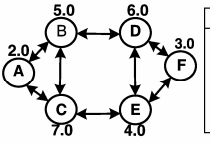
Definition 1 (*Weighed Directed Graph*)**.** A *WDG* (*W*eighted *D*irected *G*raph), denotes as *G*, is a finite set of vertices and edges, in which each edge joins one ordered pair of vertices, and each vertex or edge is associated with a weight value.

From [10], we know there exist two equivalent WDGs-- *VWDG* (*V*ertex-*WDG*) and *EWDG* (*E*dge-*WDG*). Here, we only study *VWDG* in this paper. Fig.1 (a) is *VWDG G* which has 6 vertices and 8 edges.

Definition 2 (*Traversal on Graph*)**.** A traversal on graph is a sequence of consecutive vertices along a sequence of edges on a *G*.

To easily consider, we assume each traversal has not repeated vertices. The length of a traversal is the number of vertices in it. For the Web browsing example, there are some following supposes: We consider a traversal in a small time period *t* as a traversal event. Then given a time window *l* ($l > t$), there must exist a series of traversal

events in *l*, and each vertex item in a traversal is consecutive and ordered. We also consider there is not order among these traversal events. Thus, based on the above supposes, we can regard a traversal events sequence as a traversal sequence. Each traversal event in *l* is an element of a traversal sequence. A traversal transaction database, denoted as *TTDB*, is a set of traversal sequence transactions, also denoted as *SDB*. Fig. 1(b) is a *SDB* gotten from the *VWDG* in Fig. 1(a).The length of a traversal sequence is the sum of length of all traversals contained in it. To easily consider, we assume that here the *SDB* is one individual *SDB* in which all traversal sequences come from the same an individual.



(a) Example of VWDG

TID	Traversal sequences
1	<A (ABC) (AC) D (EC) >
2	< (BC) A (EFD) (BDF) >
3	< (AC) (EDB) (DF) CB>
4	< A (CE) (DB) (EF) >

(b) A SDB from VWDG

Fig. 1. Example of VWDG and traversal sequence database D

Definition 3 (*Sup_count & Support*)**.** The support count of a sequence *S* in a *SDB D*, denoted as *sup_count* (*S*), is the number of traversal sequence containing the sequence *S*. The support of *S*, *supp* (*S*), is given by: (|*D*| be the number of traversals.)

$$supp(S) = sup_count(S) \big/ |D| \ . \tag{1}$$

Definition 4 (*Weighted Traversal Sequence*)**.** A weighted traversal sequence is a traversal sequence in which each vertex item has a weight.

The weight of the weighted traversal sequence is an average value of weights of all items in it. Given a weighted traversal sequence $S=<s_1 s_2 \ldots s_m>$, $s_j = (i_{j_1} i_{j_2} \ldots i_{j_l})$, $j \in \{1 \ldots m\}$, $l=|s_j|$, the weight of i_{j_k}, $k \in \{1 \ldots |s_j|\}$, denoted as $w(i_{j_k})$, then the weight of *S* is represented as:

$$weight(S) = \sum_{j=1}^{m} \sum_{l=1}^{|s_j|} w(i_{j_l}) \Big/ l(S) = \sum_{j=1}^{m} \sum_{l=1}^{|s_j|} w(i_{j_l}) \Big/ \sum_{j=1}^{m} |s_j| \ . \tag{2}$$

Definition 5 (*Weighted support*)**.** The weighted support of a sequence *S*, denoted as *wsupp* (*S*), is defined as follows:

$$wsupp(S) = weight(S) * (supp(S)) \ . \tag{3}$$

Definition 6 (*Weighted Frequent Traversal Sequence*)**.** A traversal sequence *S* is said to be a weighted frequent traversal sequence pattern (*WFTSP*) or weighted traversal-based sequential pattern (*WTSP*) when its weighted support is no less than a predefined minimum weighted support threshold called *wminsup*, i.e.,

$$wsupp(S) \geq minwsup \ . \tag{4}$$

Definition 7 (*Weighted Prefix*)**.** Given a weighted sequence $\alpha = <e_1e_2\ldots e_n>$ (where each e_i corresponds to a weighted frequent element in D), a sequence $\beta = <e'_1e'_2\ldots e'_m>$ ($m<n$) is called a prefix of the sequence α if (1) $e'_i=e_i$ for ($i \le m$-1), (2) $e'_m \subseteq e_m$ and (3) all the weighted frequent items in (e_m- e'_m) are alphabetically listed after those in e'_m.

Definition 8 (*Weighted Suffix*)**.** Given a weighted sequence $\alpha = <e_1e_2\ldots e_n>$. Let $\beta = <e'_1e'_2\ldots e'_m>$ ($m<n$) be the weighted prefix of α. Sequence $\gamma = < e''_m\ e_{m+1}\ldots e_n >$ is called a weighted suffix of α with regards to β, where $e''_m = (e_m - e'_m)$. If e''_m is not empty, the suffix is also denoted as <(_items in e''_m) $e_{m+1}\ldots e_n$>. If β is not a subsequence of α, the suffix of α with regards to β is empty.

Definition 9 (*Weighted Projected Database*)**.** Given a weighted sequential pattern α in a sequence database D. The α-projected database, denoted as $D|_\alpha$, is the collection of weighted suffixes of sequences in D about the prefix α.

Given a *WDG G* and a traversal *SDB D*, in this paper, our task is to find all traversal sequential patterns with weight constraint in *D*. However, the weight constraint is not an anti-monotone constraint. So we cannot directly use the '*anti-monotone*' property to prune weighted infrequent sequence patterns.

2.2 Revised Weighted Support

Suppose the weights of each item in $I = \{i_1, i_2\ldots i_n\}$ is denoted as $w(i_j)$ (j=1,2…n), it satisfy $\min(W) \le w(i_j) \le \max(W)$, where $W=\{w_1,w_2,\ldots,w_n\}$. To let the weighted support satisfy '*downward closure property*', we revise $w(i_i)$ as two following representations: $w(i_j)=min(W)$ or $w(i_j)=max(W)$. Correspondingly, the weight of *WTSP S* is revised as follows:

$$weight(S) = \sum_{j=1}^{m}\sum_{l=1}^{|s_j|} w(i_{j_l}) \Big/ l(S) = \sum_{j=1}^{m}\sum_{l=1}^{|s_j|} min(W) \Big/ \sum_{j=1}^{m}|s_j|. \quad (5)$$

or

$$weight(S) = \sum_{j=1}^{m}\sum_{l=1}^{|s_j|} w(i_{j_l}) \Big/ l(S) = \sum_{j=1}^{m}\sum_{l=1}^{|s_j|} max(W) \Big/ \sum_{j=1}^{m}|s_j|. \quad (6)$$

Thus, the '*anti-monotone property*' can be used in mining *WFTPs* since *wsupp* (S') ≤ w*supp* (S). However, if we adopt (5), we could prune some patterns which should have been weighted frequent to lead to incorrect mining results since we evaluate a lower the weight of pattern. To avoid this flaw, we adopt (6) to compute revised weight of the traversal sequence pattern S. However, the weighted support value computed by it is only an approximate value, so we should check if the pattern is really a *WTSP* with his real weight value in the last mining step.

3 Mining Weighted Sequential Patterns from WDG Traversals

Based on the fact that revised weighted setting has an '*anti-monotone*' property, we devise an efficient and scalable algorithm, called *WTSPMiner* which exploits

a '*divide-and-conquer*' strategy with an improved weighted prefix-projected pattern growth method to efficiently mine fewer but important weighted sequential patterns. *WTSPMiner* algorithm is given in Fig. 2. Figure 3 gives the procedure *WTSPM* ($SD|_{\alpha}$, *WTSP*, α, *l*), in it, *WTSP* is used to store so far found weighted traversal-based sequential patterns.

WTSPMiner(*W*eighted *T*raversal-based *S*equential *P*atterns *Miner*)

Inputs: (1)A traversal sequence transaction database ***SDB*** ; (2) A minimum support threshold: ***minwsup***, and (3) Weights of each node on graph ***G***
Output: Set of real weighted frequent traversal sequence patterns on ***G***
Method:
1. Scan *SDB* once to find the approximate weighted frequent item *b* which satisfies *supp*(*b*)**Max(W)*≥ ***minwsup***.
2. Remove all weighted infrequent item *c* which satisfies *supp*(*c*)**Max(W)*< ***Minwsup*** from each traversal sequence and sort the remained items in an element of each traversal sequence transaction alphabetically.
3. Initialize ***WTSP***. //***WTSP*** is the result set of *WFTSP*s
4. Call ***WTSPM* (*SDB, WTSP, <b>*, 1).**

Fig. 2. Algorithm WTSPMiner

***Procedure WTSPM* ($SD|_{\alpha}$, *WTSP*, α, *l*)**

Inputs: (1) $SD|\alpha$ Is the α-projected database if $\alpha \neq <>$, otherwise, it is ***SDB***. (2) ***WTSP***, the result set of weighted sequential patterns. (3) α is a real weighted traversal-based sequential patterns and (4) ***l*** is the size of α, i.e., $|\alpha|$
Output: Set of real weighted traversal-based sequential patterns
Method:
1: Scan $SD|\alpha$ once, find each candidate weighted frequent item β which satisfy *supp*(<β >)**Max*(*W*)≥ ***minwsup***, such that
   (a) β can be assembled to the last element of α to form a sequential pattern or
   (b) <β> can be appended to α to form a sequential pattern.
2: **for** ($\forall \beta \in SD|\alpha$) {
3: append β to α to form a sequential pattern α';
4: ***WTSP***=***WTSP***∪<α'>.
5: } //end of 'for' in line 2
6: **for** ($\forall S \in WTSP$) {
7: **if** ($\sum_{j=1,1}\sum_{i=1,|sj|} w(i_{ij})/l(S)$**supp*(*S*)<***minwsup***)
8: *WTSP*=*WTSP-S*; //prune the weighted infrequent traversal sequence patterns
9: **else** { // check if *S* is included in original *SDB*
10: *WTSP*=*WTSP-S*;
11: *ST*={*s*|*s* is a certain presentation of *S* & each item in element of s is the original order in original *SD*};
12: **for** ($\forall s \in ST$) {
13: **if** ($\sum_{i=1,1}\sum_{i=1,|sj|} w(i_{ij})/l(s)$**supp*(**s**) > ***minwsup***)
14: *WTSP*=*WTSP*∪*S*; // if S is a real *WTSP* then insert it to *WTSP*
15: }// end of 'for' in line 12
16: } //end of 'else' in line 9
17: } //end of 'for' in line 6
18: **for** ($\forall \alpha' \in SD|\alpha$) {
19: construct α'-projected database $SD|\alpha'$;
20: call ***WTSPM*(*SD*|α', *WTSP*, α', *l*+1)**;
21: } //end of 'for' in line 18
22: output ***WTSP***;

Fig. 3. Procedure WTSPM

Table 1. Each Prefix-projected databases and according sequential patterns

Prefix	Projected (suffix) database	Sequential patterns (minwsup=3.0)
<A>:4	<(ABC)(AC)D(CE)>,<(DEF)(BDF)>, <(_C)(BDE)(DF)CB>,<(CE)(BD)(EF)>	<AB>:4, <AC>:3,<AD>:4,<AE>:4,<AED>:3
<B>:4	<(_C)(AC)D(CE)>,<(_C)A(DEF)(BDF)>, <(_DE) (DF)CB>,<(_D)(EF)>	<B>:4,<BC>:2,<BD>:3,<BDC>:2,<BE>:3,<BF>:3, <(BC)>:2,<(BC)D>:2,
<C>:4	<(AC)D(CE)>,<A(DEF)(BDF)>, <(BDE)(DF)CB>,<(_E)(BD)(EF)>	<C>:4, <CB>:3,<C(DB)>:2,<CC>:2, <CD>:4,<CDB>:2, <CDC>:2, <CDD>:2, <CDF>:3<CE>:4,<CF>:3
<D>:4	<(CE)>,<(_EF)(BDF)>,<(_E)(DF)CB>,<(EF)>	<D>:4, <DC>:2,<DD>:2,<DF>:3
<E>:4	<(_F)(BDF)>, <(DF)CB>,<(BD)(EF)>	<E>:4,<EB>:3,<ED>:3
<F>:3	<(BDF)>,<CB>	∅

Given *minwsup*=3.0, we mine the weighted traversal sequence database shown in Fig.1 by the algorithm *WTSPMiner*. Due to the space limitation, here we do not give the detail of mining process. The resulting *WTSP*s are shown in Table 1**.**

4 Experimental Evaluations

From [6], we know that algorithm *SPAM* is by far the fastest algorithm when mining the whole set of the sequential patterns for large datasets. So we will only test the performance of *WTSPMiner* in comparison with *SPAM*. Because there are not real datasets about *WDG* currently, we test the algorithm performance using synthetic dataset. The experiments were performed on Windows XP Professional operating system with Pentium IV PC at 2.93 GHz and 768MB of RAM. We implemented our algorithm with C++ language, running under Microsoft VC++ 6.0. All the reported runtimes are in seconds.

Table 2. Experimental parameter for generating WDG and traversal sequences on it

Parameter	Meaning or parameter	Value in experiment
Vn	Number of nodes per set (i.e. items)	300
Emax	Max number of edges per vertex	1≤ Emax ≤10
wi	Weight of vertex	[0,0.33], [0.34,0.67], [0..68,0.99]
|SD|	Number of traversals in SDB	2000 to 10000
num_WSet	Number of weight set	3
Max_L	Max length of transaction	10
T	Average number of items per transaction	2.5, 9
D	Number of time windows per dataset	2000, 5000
C	Average number of elements per windows	3, 4
S	Maximum size of sequences per dataset	5, 8

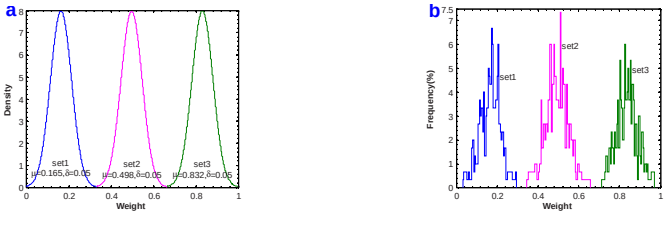


Fig. 4. (a). Gauss distribution density. (b). Distribution of weight indifferent intervals.

4.1 Generate Synthetic Datasets

In the experiment, the *DG* is generated mainly according to following parameters: number of vertices and max number of edges per vertex. And then, we randomly assigned weight to each vertex in *DG* to get the *WDG*. The characteristics of experimental parameters show in Table 2. To generate a individual *SDB*, first we generate a set *D* which includes a large number of traversal on *G*, and each traversal include *T* items, then by the random combination of the traversals, we get the *SDB*. For example, to generate sequence dataset D2KC2T3S4, we randomly select 4000 traversals from *D* whose *T*=3, then randomly divide them into 2000 partitions (can be divided

across) each which concludes 4 traversals at most. Because we study the individual *SDB* in this paper, so we assign the same weight set to the same dataset. However, we test the impact of the weight in the different value intervals to the same dataset. We assign 3 sets of weight which belongs to different range to each database respectively. The distribution of all weight sets obey Gauss distribution shown as Fig.4.

4.2 Experimental Results

In the experiments, we focused on testing the efficiency and scalability of *WTSPMiner*. In efficiency testing experiments, we compared the performance of *WTSPMiner* with *SPAM* for different datasets: one is D2C3T2.5S5 dataset and the other is D5C4T7S8 which is denser than the fist one. The performance experiment results are shown in Fig.5(a) and (b), where *minimum-sup* is a user-specified support threshold (Given a sequence *P*, *WTSPMiner* decide if *P* is frequent by *weight*(*P*)**supp*(*P*) ≥ *minimum-sup*, and *SPAM* check P is frequent by *supp*(P) ≥ *minimum-sup*.). Clearly, we can see: 1). *WTSPMiner* is faster than *SPAM* for both dataset. 2). With the decrease of *minimum-sup*, the runtime both increased and the performance difference between them becomes larger. This is because *WTSPMiner* is a weight constraint-based sequential patterns mining algorithm, and it can mine more important and has fewer search space than *SPAM* which is not constraint-based one. In addition, we can see that the weight is larger, runtime is longer. This is because that for the same *minimum-sup*, the weight is lager and then the support of pattern is smaller, i.e., the search space is larger, so the runtime is longer.

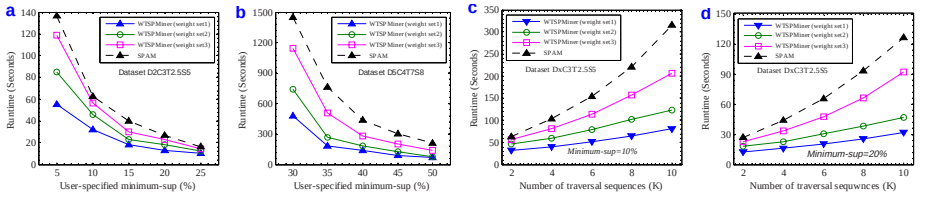


Fig. 5. Experimental results about performance and scalability of WTSPMiner and SPAM

We used DxC3T2.5S5 dataset to test algorithms' scalability with the number of traversal sequences. We range the number of traversal sequences from 2000 to 10000 in the test experiments. Figure 5(c) and (d) shows the scalability experimental results. From these results we can see: 1). Both *WTSPMiner* and *SPAM* approximately scales linearly with the size of the number of traversal sequences. 2). *WTSPMiner* has a better scalability than *SPAM*, and 3). Although itself runtime also increases, *WTSPMiner* runs faster than *SPAM* as the number traversal sequences increases. The reason for that is also *WTSPMiner* has a weight constraint to reduce the search space, but *SPAM* has not.

5 Conclusions

In this paper, we present a new algorithm, named *WTSPMiner* (Weighted Traversal-based Sequential Patterns Miner), to mine sequential patterns based on *WDG*. The key

idea of it is to take weight constraint into consideration in the mining process so as to efficiently discover fewer but more important weighted traversal patterns. *WTSPMiner* exploits an improved weighted prefix-projected pattern growth approach to decompose the task of mining original sequence database into a series of smaller tasks of mining locally projected database. To our knowledge, this is the first work to mine sequential patterns with weight-constraint from traversals on *WDG*. A comprehensive performance studies show that *WTSPMiner* is efficient and scalable for finding weighted frequent sequential patterns from traversals on *WDG*. Moreover, it can be extended to mining other kinds of frequent sequence patterns, such as substructures.

References

1. Chen, M.S., Park, J.S., Yu, P.S.: Efficient Data Mining For Path Traversal Patterns. IEEE Trans. on Knowledge and Data Engineering 10(2), 209–221 (1998)
2. Nanopoulos, A., Manolopoulos, Y.: Mining Patterns from Graph Traversals. Data & Knowledge Engineering 37(3), 243–266 (2001)
3. Agrawal, R., Srikant, R.: Mining Sequential Patterns. In: Proceedings of the 11th International Conference on Data Engineering, pp. 3–14. IEEE Press, New York (1995)
4. Han, J., Pei, J., Asi, B.M., et al.: FreeSpan: Frequent Pattern-Projected Sequential Pattern Mining. In: Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 355–359. ACM Press, New York (2000)
5. Pei, J., Han, J., Mortazavi-Asl, et al.: PrefixSpan: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth. In: Proceedings of the 17th International Conference on Data Engineering, pp. 215–224. IEEE Press, New York (2001)
6. Ayres, J., Gehrke, J.E., Yiu, T., Flannick, J.: Sequential Pattern Mining Using Bitmaps Representation. In: Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 429–435. ACM Press, New York (2002)
7. Srikant, R., Agrawal, R.: Mining sequential patterns: Generalizations and Performance Improvements. In: Apers, P.M.G., Bouzeghoub, M., Gardarin, G. (eds.) EDBT 1996. LNCS, vol. 1057, pp. 3–17. Springer, Heidelberg (1996)
8. Wang, W., Yang, J., Yu, P.S.: Efficient Mining of Weighted Association Rules (WAR). In: Proc.s of the 6th ACM SIGKDD 2000, pp. 270–274. ACM Press, New York (2000)
9. Cai, C.H., Fu, A.W.C., Cheng, C.H., et al.: Mining Association Rules with Weighted Items. In: Proc. of IDEAS 1998, pp. 68–77. IEEE Computer Society, Washington (1998)
10. Geng, R., Dong, X., Zhang, P., Xu, W.: WTMaxMiner: Efficient Mining of Maximal Frequent Patterns Based on Weighted Directed Graph Traversals. In: Proceedings of the 3rd IEEE International Conference on Cybernetics and Intelligent Systems. IEEE Computer Society, Washington (in press, 2008)

Reliable Probabilistic Classification and Its Application to Internet Traffic

Mikhail Dashevskiy and Zhiyuan Luo

Computer Learning Research Centre
Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK
zhiyuan@cs.rhul.ac.uk

Abstract. Many machine learning algorithms have been used to classify network traffic flows with good performance, but without information about the reliability in classifications. In this paper, we present a recently developed algorithmic framework, namely the Venn Probability Machine, for making reliable decisions under uncertainty. Experiments on publicly available real traffic datasets show the algorithmic framework works well. Comparison is also made to the published results.

1 Introduction

Classification and identification of network applications responsible for the creation of traffic flows is important for network management tasks such as monitoring trends of the applications in operational networks and effective network planning and design. Different learning and prediction approaches [6,8,9], including conventional statistical methods and machine learning approaches such as neural networks [1], have been applied to network traffic classification problems [10]. Despite the reported success of these methods [2,8] the learning and prediction techniques used can only provide simple predictions, i.e. the algorithms make predictions without saying how reliable these predictions are.

For learning and prediction algorithms, if we make no prior assumptions about the probability distribution of the data, other than that it is Identically and Independently Distributed (IID), there is no formal confidence relationship between the accuracy of the prediction made with the test data and the prediction associated with a new and unknown case. However, not knowing the confidence of predictions makes it difficult to measure and control the risk of error using a decision rule. In this paper, we focus on prediction with confidence and discuss one way of making reliable probability forecasting. We describe a recent developed algorithmic framework, namely the Venn Probability Machine and apply it to real traffic datasets.

2 Venn Probability Machines

We are interested in making probabilistic predictions about a sequence of examples z_1, z_2, ..., z_n where each example z_i consists of an object x_i and its label

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 380–388, 2008.
© Springer-Verlag Berlin Heidelberg 2008

y_i, $z_i = (x_i, y_i)$. The objects are elements of an object space X and the labels are elements of a finite label space Y. For example, in traffic classification, x_i are multi-dimensional traffic flow statistics vectors and y_i are application groups taking only finite number of values such as {WWW, EMAIL, P2P}. It is assumed that the example sequence is generated according to an IID unknown probability distribution P in Z^∞. The learner Γ is a function on a finite sample of n training examples $z_1, z_2, ..., z_n \in Z_n$ that makes a prediction for a new object $x_{n+1} \in X$, $\Gamma : \ Z^n \times X \to [0,1]^{|Y|}$. For each new object x_{n+1} (with true label y_{n+1} withheld from the learner), a set of predicted conditional probabilities for each possible label are produced.

The problem of making effective probability forecasts is a well studied problem in statistics, meteorology and psychology. Two simple criteria for describing how effective probability forecasts are:

1. **Reliability**: The probability forecasts *should not lie*. When a probability p is assigned to an event, there should be roughly p relative frequency of the event occurring. This is also referred to as being *well-calibrated*.
2. **Resolution**: The probability forecasts should be practically useful and enable the observer to easily rank the events in order of their likelihood of occurring. This criterion is more related to classification accuracy.

The basic idea behind the Venn Probability Machine (VPM) is to try each new test example with all possible labels and divide all the examples including the new test example into categories [5]. Under the assumption of IID (i.e. the partition of examples into groups is independent of the order that examples are presented), the construction of the VPM depends on 1) the definition of categories to cluster examples into, and 2) a mechanism of assigning individual example to these categories. The intuition behind this is that a reasonable statistical forecast would take into account only objects which are similar to the object of interest to obtain reliable estimates. The set of examples (a training set and the test example) is divided into several groups according to a criterion (which we will call *partitioning rule* or *taxonomy*). The partitioning rule should reflect similarity of the examples and are usually implemented on top of existing learning algorithm such as kNN and SVM. For example, there is no sense to put all the examples in the same group. Let X be the example space and Y the label space (the space of the classes) and $Z = X \times Y$. And let N be a predefined number of examples used as the input by the VPM. *Venn taxonomy* is a sequence A_n, $n = 1, \ldots, N$, where each A_n is a finite partition of the space $Z^{(n-1)} \times Z$. The expression $Z^{(n-1)} \times Z$ here means that A_n does not depend on the order of the first $n-1$ arguments. Let $A_n(\omega)$ be the element of the partition (A_n) that contains $\omega \in Z^{(n-1)} \times Z$. With every taxonomy (A_n) we associate a VPM. We consider a class $y \in Y$ and partition $\{(x_1, y_1), \ldots, (x_n, y)\}$ into categories, assigning two points in Z to the same category if and only if

$$A_n(\{z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n\}, z_i) = A_n(\{z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_n\}, z_j)$$

where $z_i = (x_i, y_i)$, $i = 1, \ldots, n-1$ and $z_n = (x_n, y)$. In this case we assume that $\{x_1, \ldots, z_k\}$ means a multiset (a bag), i.e. a set of elements, where each element

has a multiplicity, i.e. a natural number indicating how many memberships it has in the multiset.

For example, a simple way to define a group is based on the 1-nearest neighbour algorithm. Two examples are assigned to the same group if their nearest neighbours have the same label. This will produce a $|Y| \times |Y|$ Venn probability matrix F. Then the relative frequencies of the labels in the group containing the new example can be calculated as probabilities for the new object's label. The category T contains new example $z_n = (x_n, y)$ is nonempty and the empirical probability distribution of the label y in this category T is

$$p_y = \frac{|\{(x*, y*) \in T : y* = y\}|}{|T|}, y \in Y.$$

This is a probability distribution on Y. The VPM determined by the taxonomy is the multiprobability predictor $P_n := \{p_y : y \in Y\}$. The set P_n consists of between one and $|Y|$ distinct probability distributions on Y. The Venn Probability Machine is presented in Algorithm 1. The computational complexity of the VPM depends largely on the underlying learning algorithm for partitioning examples into categories and the label space.

A useful property of VPM is its self-calibration nature in the online learning setting. It has been proved that the probability bounds generated by the VPM is well-calibrated [5]. They bound the true conditional probability for each new test object in an online test. Using the VPM's upper and lower bounds for conditional probabilities, we can estimate the bounds for the number of errors made.

Algorithm 1. Venn Probability Machine

Require: training examples $\{(x_1, y_1), (x_2, y_2), \ldots, (x_{n-1}, y_{n-1})\}$, label space Y, new example x_n, taxonomy A_n

for $y = 1$ to $|Y|$ **do**

combine (x_n, y) with the training examples {We now have N examples $\{(x_1, y_1), ..., (x_{n-1}, y_{n-1}), (x_n, y)\}$}

partition the N examples into categories T, assigning (x_i, y_i) and (x_j, y_j) to the same category if and only if $A_n(\{(x_1, y_1), ..., (x_{i-1}, y_{i-1}), (x_{i+1}, y_{i+1}), ..., (x_n, y)\}, (x_i, y_i)) = A_n(\{(x_1, y_1), ..., (x_{j-1}, y_{j-1}), (x_{j+1}, y_{j+1}), ..., (x_n, y)\}, (x_j, y_j))$

for $i = 1$ to $|Y|$ **do**

$F_{y,i} = \frac{|\{(x*, y*) \in T : y* = y\}|}{|T|}$ {The category T containing (x_n, y) is nonempty, because it has at least this one element}

end for

end for

$Q = \arg\min_y F_{y,i}$

$i_{best} = \arg\max_i Q_i$

predict $\hat{y} = \arg\max_y F_{y,i_{best}}$

output predicted probability interval for $\hat{y}$ as $[\min_y F_{y,\hat{y}}, \max_y F_{y,\hat{y}}]$

output error probability interval [1-$\max_y F_{y,\hat{y}}$,1-$\min_y F_{y,\hat{y}}$]

output predicted mean probability for $\hat{y}^t$ as mean$_y F_{y,\hat{y}^t}$

For the implementation of VPMs, performance depends on the design of the taxonomy. The design goal is to achieve good performance in terms of test accuracy. Interestingly, the calibration property of VPMs implies that we can expect probabilistic forecasts to be calibrated, regardless of our choice of taxonomy. We often design taxonomy on the basis of existing learning algorithms that we already know perform well with respect to the class of problems we are interested in.

In this paper, we use k-Nearest Neighbours (kNN) as the basis of a taxonomy as kNN is well studied in both empirical and theoretical machine learning research. We propose two examples of taxonomy, each of them is based on the kNN algorithm. First example of taxonomy is termed "all info" where a category is formed by a set of examples which all have the same label distribution among their k nearest neighbours. The second example of taxonomy is named "voting" where a category consists of a set of examples which have the same number of most common labels among their k nearest neighbours. These two types of taxonomy give the same results when $k = 1$.

3 Experiments

We define experimental environments in which the learning tasks can be evaluated. Offline learning is the traditionally methodology of experimental testing in machine learning studies where both the training and testing sets are presented as a batch to the classifier. Online learning is a distinct learning environment where the learning algorithm is presented with unlabelled examples in sequence. It is informed of the true label of the example after it has made its prediction, but before the next unlabelled example is presented. With online learning the training dataset is continually updated as new labelled data is provided. One drawback of the online learning is that it is extremely computationally intensive. However, much of this computational burden can be avoided if the classifier generates an incremental hypothesis. The kNN classifier is an example of such a classifier, because no effective hypothesis is stored, therefore nothing needs to be updated when the new example is eventually added to the training set. The number of nearest neighbours k is one of the main parameters of a VPM implemented on kNN algorithm. In our experiment k was varied between 1 and 10.

The datasets used for our experiments are publicly available from University of Cambridge, UK at **http://www.cl.cam.ac.uk/research/srg/netos/nprobe/data/papers/sigmetrics/index.html**. The datasets have been analysed and the results have been reported in [3,9]. The data was captured from a research site (about 1000 users) connected to the Internet for a 24-hour period. Ten separate datasets ("e01"–"e10") were created based on the collected data. Due to the variation in activity during the day, there are different number of flows in each dataset. In addition, dataset "e12" was obtained at the same site some 12 month later. All datasets only contain TCP flows and there are 248 features describing simple statistics such packet length and information derived from the TCP protocol, for details see [4]. Eight classes of Internet traffic are identified {WWW, EMAIL, FTP, ATTACK, P2P, DATABASE, MULTIMEDIA, SERVICES}.

For feature selection, we utilise the filter approach as it is fast and apply the Fast Correlation-Based Filter (FCBF) using symmetrical uncertainty as correlation measure. The FCBF method has been shown effective in removing both irrelevant and redundant features [7]. By ranking all features in order of importance as computed by the method, we can simply take the top n as the most discriminate feature sets. We use the dataset "e01" as the training set to perform feature selection and the top ten features are selected for the classification experiments.

4 Empirical Results

The number of nearest neighbours k can impact the performance of the classification. The empirical results presented in Table 1 suggest that the choice of k is not critical to the performance of the kNN classifier on these datasets. These results (classification accuracy) are obtained by using the dataset "e01" as the training set and the remaining datasets as the test set. The performance comparison between kNN and kNN with feature selection is also shown in Table 1. These results confirm that feature selection can improve the performance of machine learning algorithms. In particular, with feature selection, the performance on the dataset "e12" has improved from about 66% to over 96% - a noticeable

Table 1. Performance comparison between kNN and kNN with feature selection (accuracy %)

Dataset	e02	e03	e04	e05	e06	e07	e08	e09	e10	e12
1NN	96.60	91.37	97.42	97.34	98.34	94.68	94.70	93.83	92.22	66.43
1NN FS	96.69	99.05	98.59	99.06	98.34	96.19	96.36	94.89	92.75	96.78
2NN	96.93	92.17	97.72	97.41	98.50	96.53	96.34	95.38	93.30	67.86
2NN FS	96.79	99.18	98.81	99.01	98.48	96.43	96.73	95.21	94.88	96.41
3NN	96.91	90.77	97.49	97.28	98.41	95.40	95.26	94.43	92.38	63.86
3NN FS	96.86	99.14	98.80	98.99	98.42	96.45	96.65	95.12	95.71	96.29
4NN	96.93	93.30	97.61	97.23	98.26	96.18	96.44	95.37	93.78	64.39
4NN FS	98.87	99.08	98.98	98.44	99.23	97.90	98.02	96.68	97.81	96.31
5NN	96.83	90.93	97.55	97.11	98.25	95.47	96.78	94.89	92.95	62.94
5NN FS	98.88	98.97	98.91	98.40	99.18	98.00	98.02	96.69	97.12	96.76
6NN	96.77	90.78	97.50	97.08	98.18	96.13	96.63	95.21	93.40	63.65
6NN FS	98.88	98.96	98.91	98.06	99.20	98.00	98.06	96.75	97.85	96.52
7NN	96.64	90.45	96.82	96.87	98.12	95.70	96.11	94.91	92.48	63.76
7NN FS	98.90	98.94	98.88	98.03	99.18	98.01	98.03	96.71	97.84	97.00
8NN	96.59	92.33	97.21	96.82	98.07	96.17	96.81	95.22	93.48	64.00
8NN FS	98.85	98.92	98.90	97.93	99.15	97.98	98.01	96.69	97.85	96.83
9NN	96.11	93.88	96.78	96.78	98.03	95.81	96.57	94.96	93.08	63.84
9NN FS	98.86	98.89	98.88	97.89	99.07	98.00	97.91	96.71	97.82	96.85
10NN	96.17	94.62	97.11	96.72	97.94	96.06	96.64	95.15	93.32	64.76
10NN FS	98.82	98.87	98.88	97.89	99.05	97.98	97.82	96.69	97.81	96.83

Table 2. Performance of Venn Probability Machines (accuracy %)

Dataset	e02	e03	e04	e05	e06	e07	e08	e09	e10	e12
1NN(A)	96.75	93.32	98.45	98.02	98.04	96.93	96.48	96.33	95.49	95.37
1NN(V)	96.75	93.32	98.45	98.02	98.04	96.93	96.48	96.33	95.49	95.37
2NN(A)	96.91	93.48	98.66	98.21	98.24	97.32	96.79	96.66	95.79	96.56
2NN(V)	96.74	93.32	98.42	98.02	98.04	96.93	96.46	96.31	95.48	95.39
3NN(A)	96.91	93.52	98.66	98.13	98.26	97.35	96.88	96.74	95.84	96.86
3NN(V)	96.82	91.42.	98.61	97.88	98.25	96.98	96.37	95.41	93.60	97.37
4NN(A)	96.95	93.49	98.64	98.07	98.25	97.32	96.91	96.70	94.31	97.24
4NN(V)	96.80	91.40	98.62	97.84	98.26	96.89	96.41	95.32	93.55	97.40
5NN(A)	96.94	93.42	98.65	98.08	98.21	97.33	96.89	96.70	94.35	97.35
5NN(V)	96.66	91.29	98.44	97.70	98.18	96.83	96.42	94.24	93.50	97.32
6NN(A)	96.94	93.39	98.65	98.06	98.20	97.32	96.90	96.67	94.34	97.45
6NN(V)	96.65	91.23	98.43	97.69	98.17	96.84	96.45	95.22	93.48	97.34
7NN(A)	96.83	93.42	98.68	98.08	98.20	97.32	96.88	96.69	94.34	97.48
7NN(V)	96.60	91.05	98.38	97.56	98.17	96.82	96.43	95.16	93.39	97.28
8NN(A)	96.81	93.39	98.64	98.07	98.20	97.31	96.85	96.67	94.37	97.50
8NN(V)	96.60	91.06	98.42	97.57	98.17	96.82	96.43	95.18	93.40	97.29
9NN(A)	96.82	93.38	98.67	98.06	98.23	97.28	96.87	96.68	94.37	97.57
9NN(V)	96.60	91.05	98.39	97.51	98.11	96.79	96.36	95.10	93.34	96.68
10NN(A)	96.80	93.40	98.62	98.08	98.27	97.27	96.87	96.67	94.39	97.57
10NN(V)	96.58	91.03	98.41	97.51	98.10	96.79	96.39	95.49	93.33	96.71
NBKE	92.21	95.91	93.98	94.37	94.94	89.41	90.19	86.78	90.34	88.14

increase of over 30% in overall classification accuracy. This result indicates that the selected features are very stable over the times. In the following experiments, we use these ten features for probability forecasting based on VPM. The same datasets have been analysed using the Naive Bayes estimator [3] and the reported performance is similar to that of kNN shown in Table 1. For example, it is reported that the best average percentage of accuracy of Naive Bayes performed on all features and on the selected features is 93.50% and 96.29%, respectively.

In the offline setting, we again use the dataset "e01" as the training set and make classification on all other 10 datasets. Two different types of taxonomy described in the Section 2 are used and the experimental results are presented in Table 2. As expected, their performance is identical when k=1. However, the performance using "all info" taxonomy (denoted as A in the table)is often better than that of using "voting" taxonomy (denoted as V in the table) as the "all info" taxonomy is able to partition the homogenous examples together. The performance of both types of taxonomy is comparable with those obtained by using conventional kNN methods except on the dataset "e03".

For online setting, we evaluate the performance of VPM implemented on the top of kNN. For illustration purpose, the results of VPM implemented on the top of 3NN tested on the dataset "e12" (randomly permuted) are presented in Fig. 1, where the following three curves are plotted against the number of examples n:

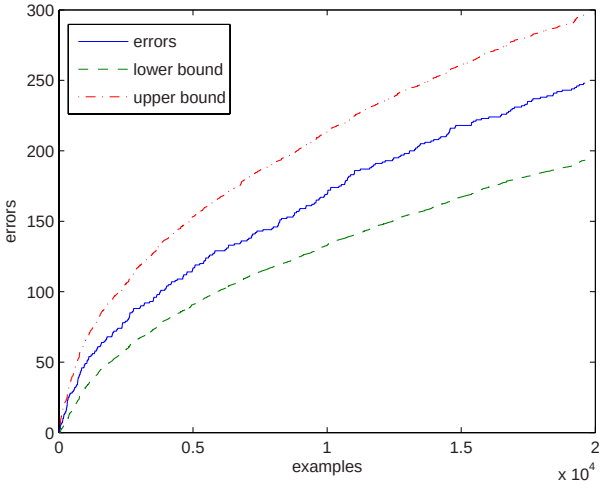


Fig. 1. Performance of the Venn Probability Machine (3NN) on the dataset "e12" in online mode

- the cumulative error curve (solid lines) $E_n = \sum_{i=1}^{n} err_i$ where err_i =1 if an error is made at trial i and otherwise $err_i = 0$.
- the cumulative lower error probability curve (dashed lines) $L_n = \sum_{i=1}^{n} l_i$
- the cumulative upper error probability curve (dotted lines) $U_n = \sum_{i=1}^{n} u_i$

where $[l_i, u_i]$ is the error probability interval output by the algorithm at trial i for the label y_i, see Algorithm 1. As shown in Fig. 1, the cumulative number of errors E_n is 248 (i.e. the mean error rate is 1.264%) and the mean probability interval is [0.999%, 1.521%]. These numbers confirm that the error probability intervals are well calibrated.

We also want to assess the quality of probability forecasts made VPM. Probability metrics are minimised when the predicted value of each example is the true underlying probability of that example being positive. Root mean squared error is an example of such probability metrics for assessing the quality of probability forecasts. We calculate root mean squared error (RMSE) of the predictions made by a classifier as $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|Y|}\sum_{j=1}^{|Y|}(\mathbb{I}_{\{y_i=j\}} - \hat{p}_{i,j})^2}$, where N is the number of examples, $|Y|$ is the number of classes, $\mathbb{I}_{\{y_i=j\}}$ is the indicator function which taking 1 if $y_i = j$, 0 otherwise, and $\hat{p}_{i,j}$ is the probabilistic forecast for the ith example with the label j.

In the offline setting, we again use the dataset "e01" as the training set and make predictions on all other 10 datasets. Table 3 shows the RMSE of Venn Probability Machines using "all info" (A) and "voting" (V) taxonomy. For comparison purpose, we also run the Naive Bayes estimator which has been used in [3] on the same datasets. Table 3 shows the performance of Naive Bayes with kernel estimation (NBKE) and these results demonstrate that Venn Probability Machines perform really well on these datasets.

Table 3. The performance of Venn Probability Machines (RMSE)

Dataset	e02	e03	e04	e05	e06	e07	e08	e09	e10	e12
1NN(A)	0.0896	0.1281	0.0616	0.0693	0.0691	0.0857	0.0922	0.0944	0.1051	0.1046
1NN(V)	0.0892	0.1281	0.0616	0.0693	0.0691	0.0857	0.0922	0.0944	0.1051	0.1046
2NN(A)	0.0867	0.1273	0.0566	0.0661	0.0644	0.0819	0.0877	0.0907	0.1038	0.0861
2NN(V)	0.0892	0.1281	0.0619	0.0693	0.0691	0.0857	0.0924	0.0945	0.1052	0.1045
3NN(A)	0.0875	0.1276	0.0568	0.0666	0.0635	0.0822	0.0862	0.0908	0.1039	0.0836
3NN(V)	0.0889	0.1458	0.0585	0.0722	0.0656	0.0862	0.0945	0.1060	0.1259	0.0796
4NN(A)	0.0880	0.1281	0.0568	0.0669	0.0637	0.0822	0.0871	0.0914	0.1073	0.0834
4NN(V)	0.0891	0.1460	0.0582	0.0728	0.0655	0.0875	0.0941	0.1071	0.1263	0.0794
5NN(A)	0.0884	0.1284	0.0576	0.0668	0.0640	0.0820	0.0871	0.0903	0.1070	0.0821
5NN(V)	0.0911	0.1468	0.0619	0.0749	0.0669	0.0884	0.0939	0.1079	0.1268	0.0805
6NN(A)	0.0884	0.1280	0.0577	0.0661	0.0642	0.0821	0.0874	0.0900	0.1066	0.0816
6NN(V)	0.0912	0.1473	0.0622	0.0752	0.0671	0.0882	0.0935	0.1081	0.1270	0.0803
7NN(A)	0.0893	0.1262	0.0581	0.0656	0.0642	0.0819	0.0873	0.0900	0.1065	0.0804
7NN(V)	0.0918	0.1486	0.0630	0.0770	0.0672	0.0883	0.0937	0.1086	0.1278	0.0812
8NN(A)	0.0897	0.1258	0.0581	0.0657	0.0644	0.0821	0.0875	0.0899	0.1065	0.0805
8NN(V)	0.0918	0.1485	0.0623	0.0770	0.0671	0.0885	0.0937	0.1084	0.1277	0.0811
9NN(A)	0.0896	0.1257	0.0580	0.0656	0.0644	0.0822	0.0874	0.0898	0.1065	0.0803
9NN(V)	0.0919	0.1486	0.0630	0.0778	0.0680	0.0888	0.0947	0.1093	0.1283	0.0895
10NN(A)	0.0897	0.1256	0.0582	0.0656	0.0637	0.0821	0.0875	0.0900	0.1065	0.0806
10NN(V)	0.0921	0.1487	0.0627	0.0780	0.0684	0.0888	0.0942	0.1093	0.1284	0.0891
NBKS	0.1289	0.0976	0.1137	0.107	0.1059	0.1476	0.1399	0.1653	0.1421	0.1633

5 Conclusions

In this paper we addressed probabilistic prediction with confidence and presented a newly developed probabilistic prediction model, namely Venn Probability Machine (VPM). The property of validity guaranteed by VPM is the probabilistic predictions output by Venn predictors will be "well calibrated". We then presented the VPM implementation based on kNN and proposed two examples of taxonomy. Finally we assessed the performance of VPM implementation on publicly available real traffic datasets. Empirical results demonstrated good performance.

We believe that the framework presented in this paper is practically useful for many applications such as network resource management. If the network provider knows that the predictions have been made by a reliable and reasonably accurate forecasters, they are then in the position to wait for the most ideal situation to act. We plan to apply this framework to the resource management for next-generation optical networks.

Acknowledgements

This work has been supported by the EPSRC through grant EP/E00053/1 "Machine Learning for Resource Management in Next-Generation Optical Networks".

References

1. Auld, T., Moore, A., Gull, S.: Bayesian Neural Networks for Internet Traffic Classification. IEEE Transactions on Neural Networks 18(1), 223–239 (2007)
2. Cosmas, J.P., Pitts, J.M., Luo, Z., Bocci, M., Nyong, D., Rai, S.: Identification of Resource Allocation Trends in Irregular ATM Networks. In: Proceeding of IFIP Workshop TC6, Fourth Workshop on Performance Modelling and Evaluation of ATM Networks (July 1996)
3. Andrew, M., Denis, Z.: Internet Traffic Classification using Bayesian Analysis Techniques. In: SIGMETRICS 2005, pp. 50–60 (2005)
4. Andrew, M., Denis, Z., Crogan, M.: Discriminators for Use in Flow-Based Classification. Technical Report RR-05-13, Dept. of Computer Science, Queen Mary, University of London, ISSN 1470-5559 (2005)
5. Vovk Vl, Alex, G., Shafer, G.: Algorithmic Learning in a Random World. Springer, Heidelberg (2005)
6. Williams, N., Zander, S., Armitage, G.: A Preliminary Performance Comparison of Five Machine Learning Algorithms for Practical IP Traffic Flow Classification. ACM SIGCOMM Computer Communication Review 36(5), 7–15 (2006)
7. Yu, L., Liu, H.: Efficient Feature Selection via Analysis of Relevance and Redundancy. Journal of Machine Learning Research 5, 1205–1224 (2004)
8. Zander, S., Nguyen, T., Armitage, G.: Automated Traffic Classification and Application Identification using Machine Learning. In: The IEEE Conference on Local Computer Networks, pp. 250–257 (2005)
9. Denis, Z., Andrew, M.: Traffic Classification Using a Statistical Approach. In: PAM, pp. 321–324 (2005)
10. Yunhao, L., Joseph, F., Ioannis, L., Huang, C.C.: Traffic Classification and service in wavelength routed all-optical networks. In: IEEE International Conference on Communications, vol. 2, pp. 1375–1380 (2003)

A New Decision Rule for Statistical Word Sense Disambiguation

Dongmei Fan, Zhimao Lu, and Rubo Zhang

Harbin Engineering University
Harbin, China
{fandongmei,lzm,zrbzrb}@hrbeu.edu.cn

Abstract. Word Sense Disambiguation (WSD) is usually considered to be a pattern classification to research and it has always being a key problem and one of difficult points in natural language processing. Statistical learning theory is a mainstream of the research method for WSD. The distribution of the word-senses of an ambiguous word is always not symmetrical and the distinction between word-senses' emergence frequency is great sometimes, so the judgment results are inclined to the maximum probability word-sense in the word-sense classification. The reflection of this phenomenon is obviously in the Bayesian model. When using the Bayesian model to carry on some research we find a new word-sense decision rule, which have a better precision than Bayesian model in WSD. In order to validate the credibility and stabilization of this method we carry through the experiment time and again, and acquire lots of experiment data. The results of the experiment indicate that new decision rule is more excellent than Bayesian decision rule. Furthermore this paper provides a theoretical foundation for this new decision rule.

1 Introduction

To determine the sense of an ambiguous word in a specific context, word sense disambiguation (WSD) has been a hot topic in natural language processing. It is an important technique for information retrieval, text mining, machine translation, text classification, automatic text summarization etc.

Stochastic solutions to WSD acquire linguistic knowledge from the training corpus by statistical methods, and apply the knowledge to disambiguation. The first stochastic model in WSD was built by Brown in 1991[1]. Since then, most machine learning methods have been applied to WSD, including decision tree, Bayesian model, neural network, SVM, maximum entropy, genetic algorithms etc[2-4].

Two main research topics in statistical WSD are the statistical method (machine learning) and knowledge source (corpus or dictionary)[5-7] For different learning methods, supervised methods usually get good performance at a cost of human tagging of training corpus. The precision improves with larger size of training corpus. Human tagging is a hard job. Unsupervised methods do not need the tagging of corpus, but it usually acquires less precision than supervised methods. Knowledge acquisition is a key to WSD methods [8-10].

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 389–396, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2 Issues Raised

2.1 Phenomenon P-1

In WSD based on the statistical learning, if ambiguous word meaning frequency (for training data,) the disparity between the time difference, meaning more inclined to judge results of the greatest probability word-sense. It is said Phenomenon P-1in this paper.

In the large-scale real corpus, the distribution of ambiguous word' senses is usually uneven, and sometimes, and some meaning when the probability distribution of difference will be very poor, such as the word-sense of some commonly used high probability, or even near 100%, while the probability of other word-senses small but can be ignored.

Gale suggests a lower bound or baseline value for WSD performance [11]. Baseline refers to the WSD precision of the simplest method, usually the percentage of the maximum frequency. For a two-sense ambiguous word, a 90% precision is its baseline if the frequency distribution is 9:1 between the two senses, but this result is very high, not easy to surpass.

The reason is that at this time the impact of the word-sense with maximum probability is becoming quite tremendous, resulting in a situation described by phenomenon P-1. This phenomenon displays more prominently when using the WSD method based on Bayesian model.

In order to discuss the above issue, this paper needs to start from the Bayesian decision rule. The mathematical expression of the Bayesian decision rules is as follows:

$$\begin{aligned} s' &= \arg\max_{s_k} P(s_k|C_{context}) \\ &= \arg\max_{s_k} \frac{P(C_{context}|s_k)}{P(C_{context})} P(s_k) \\ &= \arg\max_{s_k} P(C_{context}|s_k)\, P(s_k) \\ &= \arg\max_{s_k} [\log P(C_{context}|s_k) + \log P(s_k)] \end{aligned} \tag{1}$$

Among them, sk $\neq s'$, Ccontex is the contextual environment(i.e. a collection of context words), sk is any word-sense variable of ambiguous words, s´ is correct word-sense, P is the probability.

The Parameter P(sk) of formula(1), as the appearance probability of ambiguous words, might be the main reason for phenomenon P-1. In this experiment, this paper make the word-sense judgment in accordance with the formula(2) deliberately excluding the addition of log P (sk), and resulted in the unexpected phenomenon that the correct rate (or F-Measure) was even higher than the use of formula (1), such instance is regarded as phenomenon P-2 in this paper.

$$s' = \arg\max_{s_k} P(C_{context}|s_k) \tag{2}$$

2.2 Phenomenon P-2

Formula (2) has much higher accuracy than formula (1), for the judgment of word-sense. It is said Phenomenon P-2in this paper.

It seems that formula (2) which has not calculated P(sk) is lack of necessary theoretical basis for the reason that it does not conform to Bayesian probability formula. Why, then, would appear in the P-2? May be a calculate exception of statistical method? To explore this issue, the paper redesigns the experiment that is doing WSD experiment for 20 Chinese ambiguous words provided in Senseval-3 using the means of word-sense decision-making method and then making contrastive analysis with the method carried by Lu Zhimao in 2006[9]. The experimental results in terms of the 19 F-Measure what changes have taken place not only the results of (the correct rate and the F-Measure) of the ambiguous word (Di1 Fang1)" appeared markedly improved, see Table I. One accuracy and F-Measure reference to the calculating method the literature [9].

Table 1. Comparative experiment results of ambiguous word "地方"

The number of word-sense	Accuracy		F-Measure	
4	formula(1)	formula(2)	formula(1)	formula(2)
	0.7059	0.7647	0.6511	0.7157

There is only one experimental result which changed in the 20 ambiguous words, the rate is 1 to 19, is it really an exception? This paper carefully examines the training data provided by Senseval-3 again, and finds out that word-sense probability distribution (for training and testing corpus) of each word in the 20 ambiguous words is even, besides "地方" . That is:

$$P(s_1) = P(s_2) = \cdots = P(s_k) \tag{3}$$

Under this condition, P(sk), as the constant, causes formula (1) degenerates to formula (2) in the form, therefore the experimental results by using formula (1) and formula (2) are the same. However, ambiguous words "地方" distribution is actually imbalanced for the frequency ratio is 6: 3: 4: 5, and when using formula (2) to carry on the disambiguation experiment, accuracy would be higher, that means maybe the emergence of phenomenon seems to be not for the simple understanding of fortuitous events.

Since the phenomenon p-2 can not be easily assert to be an exception, this paper names formula (2) the new word-sense decision-making rule New Rule, NR). If it is a new word-sense judgment method, it will need to give a theoretical and experimental verification, and this is the task of this paper.

3 Experiments

In order to deeply discuss the reason of phenomenon P-2, it needs thorough and meticulous analysis and experiments to validate word-sense decision new rule (NR). The

Senseval-3 data quantity is too small that each word-sense training example is at most 20 and the testing cases are just within 10. Training data of statistical Learning requirements must be large enough; otherwise the effect will be very unstable and unreliable, then it will be difficult to judge the merits of the method.

In order to obtain a large-scale, high-quality marked corpus for comparative study of WSD, the best way is to use artificial ambiguous word[2]. Although artificial ambiguous words different from the true word ambiguity, but it can be very good simulation, simulation Disambiguation process. Meaning in particular Disambiguation method comparison, the artificial ambiguous word can play a bigger role. As we all know, the use of artificial methods of experimental corpus for Meaning tagging and processing, it will be hard to scale large enough corpus (such as the International Evaluation corpus Senseval [12]), and the Manual tagging of Corpus no good consistency . And the term is used artificial ambiguous word can be solved both problems.

The purpose of the experiment in this paper is comparing Bayesian decision-making rule(regarded as NB for short) with word-sense decision new rule(NR), so this paper take the artificial ambiguous word as the experimental Object of WSD in the follow experiments.

According to statistics, the average word-sense quantity of Chinese ambiguous word is 2.41, that is to say the word-sense quantity of vast majority ambiguous words is 2. Generally ambiguous words have several different word-sense parts of speech (POS), some of them are the same, and some different. In practice, therefore, when constructing artificial ambiguous words, it is need to choose a number of different POS content words (to be single-word) to construct 8 two-senses ambiguous words and 6 three-senses ambiguous words, furthermore, the quantity of other artificial ambiguous words is relatively less, see also Table 2. wr in the table represents

Table 2. Artificial ambiguous word structure Table

Wr	**Morpheme**	Frequency ratio of word-sense	*Wr*	**Morpheme**	Frequency ratio of word-sense
w_1	出现/避免	9:1	w_2	出现/成果	9:1
w_3	保护/优秀	8:2	w_4	保护/避免	8:2
w_5	保护/技术	7:3	w_6	举行/群众	7:3
w_7	举行/优秀	6:4	w_8	保护/举行	6:4
w_9	出现/避免/优秀	8:1:1	w_{10}	举行/成果/普通	7:2:1
w_{11}	出现/群众/重要	6:3:1	w_{12}	保护/技术/普通	6:3:1
w_{13}	保护/产生/群众	5:3:2	w_{14}	产生/优秀/成果	5:4:1
w_{15}	出现/避免/保护/成果	6:2:1.5:0.5	w_{16}	出现/保护/优秀/技术	5:2.5:1.5:1
w_{17}	举行/保护/重要/群众	4:3:2:1	w_{18}	举行/出现/优秀/成果	4:3:2:1
w_{19}	出现/保护/避免/优秀/成果	3:2.5:2:1.5:1	w_{20}	出现/产生/举行/普通/群众	4:2.5:2:1:0.5
w_{21}	举行/出现/保护/产生/优秀/群众	3:2.5:2:1.5:0.5:0.5	w_{22}	举行/出现/保护/避免/普通/技术	2.5:2.5:2:1.5:1:0.5

artificial ambiguous word, the frequency ratio of ambiguous word-sense is:

$$r_1 : \mathrm{r2} : r_3 \ \ldots \ r_n$$

$$\sum_{i}^{n} r_i = 10 \tag{4}$$

And *n* represents that the ambiguous words have *n* word-senses.

This paper has constructed 22 ambiguous words whose word-sense quantity is 2~6 and has controlled probability distribution of ambiguous words to ensure that there is a wide difference between different word-senses in the first experiment.

3.1 Experimental Result

The number of Examples of each ambiguous word was controlled within 4000, and scale-proportion of training and testing corpus was controlled within 3:1. In the experiment, Bayesian Decision rules(NB) and new rules(NR) were adopted to establish the word-sense classification model, to train and test against 20 ambiguous words.

In order to calculate P(Ccontext|sk), it is also needed to assume that it is conditionally independent between every word in the context of ambiguous words. With this assumption, P(Ccontext|sk) could be calculated by the formula of multiple multiplication, see formula (5) :

$$\begin{aligned} P\left(C_{context}|s_k\right) &= P(\{v_j \mid v_j \text{ in } C_{\text{context}}\} \mid s_k) \\ &= \prod_{v_j \text{ in } C_{\text{context}}} P\left(v_j|s_k\right) \end{aligned} \tag{5}$$

In the formula (6), v_j was the j^{th} word in context.

Because the two methods use probabilistic model, Baseline of experimental results was determined by maximum probability semantic method in this paper[2], see the experimental results in Tables 3.

Every word-sense of ambiguous words has an F-Measure, and the F-Measure of ambiguous words needs to weight the average of all the F-Measure of word-senses:

$$F_A = \sum_{i}^{n} F_i P_i \tag{6}$$

In the formula above, FA is the F-Measure of ambiguous words, Fi is the F-Measure of ith word-sense, Pi is the appearance probability of ith word-sense, n is the number of word-senses.

Accuracy is calculated in Table 3, but at this time C(correct) represents the correct number in all the tagged examples and C(tagged) is the number of all tagged examples. It is different from calculating the accuracy of some ambiguous word-sense for that the numbers of tagged results and tagged correctly results are aimed at word-sense.

3.2 Analysis of the Experimental Results

According to Table 3, it could be seen that experimental accuracy of both NB and NR both have exceeded each corresponding Baseline basically wherever word-sense number of ambiguous words is more or less. That is to say that the two methods are effective in WSD. Of course, it is not easy to exceed the Baseline for two word-senses ambiguous words w1 and w2 whose frequency ratio is 9:1, but Baseline reaches 90%.

Table 3. Experimental results

Ambiguous word	Accuracy			F-Measure	
	Baseline	NB	NR	NB	NR
w_1	0.9	0.9	0.9	0.8526	0.8526
w_2	0.9	0.9	0.9	0.8526	0.8526
w_3	0.8	0.821	0.827	0.7575	0.7695
w_4	0.8	0.8	0.802	0.7111	0.7159
w_5	0.7	0.749	0.753	0.6789	0.6862
w_6	0.7	0.956	0.957	0.9549	0.956
w_7	0.6	0.897	0.897	0.8963	0.8964
w_8	0.6	0.884	0.886	0.8807	0.8829
w_9	0.8	0.8	0.801	0.7111	0.7135
w_{10}	0.7	0.738	0.743	0.6543	0.663
w_{11}	0.6	0.745	0.756	0.69	0.7033
w_{12}	0.6	0.686	0.695	0.6134	0.6264
w_{13}	0.5	0.645	0.653	0.5913	0.6048
w_{14}	0.5	0.719	0.72	0.678	0.6792
w_{15}	0.6	0.601	0.606	0.4675	0.4775
w_{16}	0.5	0.728	0.734	0.6833	0.6899
w_{17}	0.4	0.721	0.725	0.67	0.6762
w_{18}	0.4	0.743	0.747	0.7066	0.7118
w_{19}	0.3	0.592	0.595	0.5485	0.5548
w_{20}	0.4	0.614	0.624	0.5647	0.5768
w_{21}	0.3	0.632	0.638	0.5713	0.5799
w_{22}	0.25	0.65	0.655	0.6116	0.6177

What is gratifying is that the accuracy and F-Measure of twenty in twenty two ambiguous words have upgraded after using New Rules (NR) of word-sense decision. The amplitude of F-Measure was calculated with formula (7) in this paper:

$$Amp = \frac{F_{NR} - F_{NB}}{F_{NB}} \times 100\% \tag{7}$$

In the formula above, Amp represents the amplitude of F-measure, FNR represents the F-measure of NR model, FNB represents the F-measure of NB model.

The amplitude of w13 (three-senses ambiguous word) nearly approach 2.3 percentage points is the largest. in addition, three amplitudes of F-Measure is over 2.1 percentage points, ten amplitudes is more than 1.0 percentage points, and fourteen amplitude is more than 0.68 percentage points. For the ambiguous words which word-senses are more than three, it is more difficult to be disambiguated, but the amplitude of F-measure for all the ambiguous words in this experiment increase significantly. For instance, the amplitudes of five ambiguous words w15、w19、w20、w21、w22) are more than 1.0 percentage points and other three ones (w16、w17、w18) also surpass 0.74 percentage points.

Apart from the ambiguous words w1 and w2, all other experimental results have been obviously promoted; it shows that the role of multiplication factor P (sk) is negative

when word-sense frequency difference is not a great disparity. Simultaneously it also shows that P (sk) is at least one of the main reasons of phenomenon P-1.

In this experiment, every ambiguous word has 3,000 training examples and 1,000 test examples, then, the total scale of experimental corpus for 22 ambiguous words has reached 88,000 examples which are abundant. The experimental results can prove that the phenomenon P-2 was not accidental, and NR is more accurate than NB in word-sense distinction.

4 Conclusion

NB is a classic statistical learning method which is widely used in pattern recognition and classification problems. Through the experiments, Zhimao Lu proved that the calculation of NB is more convenient and its accuracy of disambiguation is comparable to the common statistical classification methods such as the maximum entropy method, support vector machine, decision tree, BP neural network and so on[13]. If complemented by feature selection method, Bayesian model will be more outstanding in WSD. It is discovered that word-sense judgment method based on NR is more precise than NB in this paper, which is proved by statistical learning theory though the experiments. Also this paper predicted that if NR is supplemented by appropriate feature selection method, the effect of WSD will be also enhanced.

However, there are many kinds of feature selection method for word-sense classification, so it is need to be further studied about which is more suitable for NR-based model further more. Through research and experiments, it is considered that these studies will actively promote the development of statistical WSD study.

Acknowledgement. The authors would like to thank the anonymous reviewers for their helpful comments. This work is supported by the National Natural Science Foundation of P.R. China (Grant: 60603092) and the Ph.D. Programs Foundation of Ministry of Education of China (Grant: 20070217043).

References

1. Peter, F.B., Stephen, A., Della, P.: Word-Sense Disambiguation Using Statistical Methods. In: Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics, pp. 264–270 (1991)
2. Gale, W., Church, K., Yarowsky, D.: A Method for Disambiguating Word Senses in a Large Corpus. Computers and the Humanities 26, 415–439 (1992)
3. Bhattacharya, I., Getoor, L., Bengio, Y.: Unsupervised Sense Disambiguation Using Bilingual Probabilistic Models. In: Proceedings of the 42th Annual Meeting of the Association for Computational Linguistics (2004)
4. Yee, S.C., Hwee, T.N.: Estimating Class Priors in Domain Adaptation for Word Sense Disambiguation. In: Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL, pp. 89–96 (2006)
5. Mona, D.: Relieving the Data Acquisition Bottleneck in Word Sense Disambiguation. In: Proceedings of the 42th Annual Meeting of the Association for Computational Linguistics (2004)

6. Navigli, R.: Meaningful Clustering of Senses Helps Boost Word Sense Disambiguation Performance. In: Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL, pp. 105–112 (2006)
7. Xue, N.W., Chen, J.Y.: Martha Palmer. Aligning Features with Sense Distinction Dimensions. In: Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions, pp. 921–928 (2006)
8. Mona, D.: An Unsupervised Approach for Boot-strapping Arabic Word Sense Tagging. In: Proceedings of Arabic Based Script Languages (2004)
9. Lu, Z.M., Wang, H.F., Yao, J.M., Li, S.: An Equivalent Pseudoword Solution to Chinese Word Sense Disambiguation. In: Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (Coling-ACL 2006), Sydney, Australia, pp. 17–21 (2006)
10. Lyle, U., Martha, P.: An Empirical Study of the Behavior of Active Learning for Word Sense Disambiguation. In: Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL, New York, pp. 120–127 (2006)
11. William, A.G., Church, K.W., Yarowsky, D.: Estimating Upper and Lower Bounds on the Performance of Word Sense Disambiguation Programs. In: Proceedings of the 80th Annual Meeting of the Association for Computational Linguistics, Newark, Delaware (1992)
12. Ted, P.: Machine Learning with Lexical Features: The Duluth Approach to SENSEVAL-2. In: Proceedings of Senseval-2, Second International Workshop on Evaluating Word Sense Disambiguation Systems, France, July 2001, pp. 139–142. SIGLEX, Association for Computational Linguistics, Toulouse (2001)
13. Lu, Z.M.: Study on Chinese Word Sense Disambiguation Based on Unsupervised Methods. Harbin Institute of Technology. PhD Thesis, China (2006)

A Vicarious Words Method for Word Sense Discrimination

Zhimao Lu, Dongmei Fan, and Rubo Zhang

Harbin Engineering University
Harbin, China
{lzm,fandongmei,zrbzrb}@hrbeu.edu.cn

Abstract. This paper presents a new approach based on Vicarious Words (VWs) to resolve Word Sense Discrimination (WSD) in Chinese language. VWs are particular artificial ambiguous words, which can be used to realize unsupervised WSD. A Bayesian classifier is implemented to test the efficacy of the VW solution on Senseval-3 Chinese test suite. The performance is better than state-of-the-art results with an average F-measure of 0.80. The experiment verifies the value of VW for unsupervised method in WSD.

1 Introduction

To determine the sense of an ambiguous word in a specific context, word sense Discrimination (WSD) has been a hot topic in natural language processing. It is an important technique for information retrieval, text mining, machine translation, text classification, automatic text summarization etc [1].

Stochastic solutions to WSD acquire linguistic knowledge from the training corpus by statistical methods, and apply the knowledge to discrimination. The first stochastic model in WSD was built by Brown in 1991[2]. Since then, most machine learning methods have been applied to WSD, including decision tree, Bayesian model, neural network, SVM, maximum entropy, genetic algorithms etc [3].

Two main research topics in statistical WSD are the statistical method (machine learning) and knowledge source (corpus or dictionary) [4]. For different learning methods, supervised methods usually get good performance at a cost of human tagging of training corpus. The precision improves with larger size of training corpus. Human tagging is a hard job. Unsupervised methods do not need the tagging of corpus, but it usually acquires less precision than supervised methods. Knowledge acquisition is a key to WSD methods [5]. This paper aims at an unsupervised method, which acquires the WSD knowledge from raw corpus. An equivalent pseudo word solution is proposed to meet that ends.

2 Conception of Vicarious Word

Is it possible to acquire knowledge for WSD by automatic mining from raw corpus? It seems impossible without even a “seed” tagging or any parallel corpus. But even raw corpus without any human tagging has some useful information in it. Because

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 397–404, 2008.
© Springer-Verlag Berlin Heidelberg 2008

monosemous words are unambiguous priori knowledge, and explain 86%~89% instances in a dictionary and 50% items in running corpus, they are potential knowledge source for WSD.

A monosemous word is usually synonymous to some polysemous words. This is quite common in Chinese, which should be a knowledge source for WSD.

2.1 Conception of the Vicarious Word

If replacing the ambiguous words in the corpus with its synonymous monosemous word, then isn't it convenient to acquire knowledge from raw corpus? For example in table 1, the ambiguous word "把握" has three senses, whose synonymous monosemous words are listed on the right column. These synonyms contain some information for WSD.

An artificial ambiguous word can be coined with the monosemous words in table 1. This process is similar to the making of general pseudowords [6-8], but has some essential difference. Because this artificial ambiguous word need to simulate the function of the real ambiguous word, and to acquire semantic knowledge as the real ambiguous word does, we call it a vicarious word (VW) for its equivalence with the real ambiguous word. It's apparent that the Vicarious Word has provided a new way to unsupervised WSD.

Table 1. Synonymous monosemous words for the ambiguous word "把握"

把握(ba3 wo4)	S_1:	信心/自信心
	S_2:	握住/在握/把住/抓住/控制
	S_3:	领会/理解/领悟/深谙/体会

The equivalence of the VW with the real ambiguous word is a kind of semantic synonym or similarity, which demands a maximum similarity between the two words. An ambiguous word has the same number of VWs as of senses. Each VW's sense maps to a sense of ambiguous word.

The semantic equivalence demands further equivalence at each sense level. Every corresponding sense should have the maximum similarity, which is the strictest limit to design and construction of a VW.

The starting point of unsupervised WSD based on VW is that VW can substitute the original word for knowledge acquisition in model training. Every instance of each morpheme of the VW can be viewed as an instance of the ambiguous word, thus the training set can be enlarged easily. VW is a solution to data sparseness for lack of human tagging in WSD.

2.2 Basic Assumption for VW-Based Idea

It is based on the following assumptions that VWs can substitute the original ambiguous word for knowledge acquisition in WSD model training.

Assumption 1: Words of the same meaning play the same role in a language. The sense is an important attribute of a word. This is a common sense and plays as the basic assumption here.

Assumption 2: Words of the same meaning occur in similar context. This assumption is widely used in semantic analysis and plays as a basis for much related research. For example, some researchers make clustering of the context of ambiguous words for WSD, which shows good performance [9]. Other research works also prove the efficacy of this assumption in WSD solutions.

Because a VW has the same or enough similarity with the ambiguous word in syntax and semantics, it is a useful knowledge source for WSD.

2.3 Implementation to the VW-Based Solution

An appropriate machine-readable dictionary is needed for construction of the VWs. TongYiCiCilin (a Chinese thesaurus [10]) is adopted and revised to meet this demand.

According to the position of the ambiguous word, a proper word is selected as the morpheme of the VW. Nearly every ambiguous word has its corresponding VW constructed in this way.

The first step is to decide the position of the ambiguous word starting from the leaf node of the tree structure. Words in the same leaf node are identical or similar in the linguistic function and word sense. Other words in the leaf node of the ambiguous word are called brother words of it. If there is a monosemous brother word, it can be taken as a candidate morpheme for the VW. If there doesn't exist such a brother word, trace to the fourth level. If there is still no monosemous brother word in the fourth level, trace to the third level. Because every node in the third level contains many words, it satisfies the morpheme search most of the time. This is a kind of search from bottom to up.

In most cases, the synonym can be found at the fifth level. It's not often necessary to search to the fourth level, less to the third. According to our research, TongYiCiCilin(extended) contains corresponding monosemous words for 93% of the ambiguous words in the fifth level, and 97% in the fourth level. There are only 112 ambiguous words left, which counts for the other 3% and mainly are functional words. Some of the 3% words are rarely used which cannot be found in even a large running corpus. Words that lead to semantic misunderstanding are usually content words. In WSD research in English, only nouns, verbs, adjectives and adverbs are considered. From this aspect, we can see that the extended version of TongYiCiCilin (extended) meets our demand for the construction of VWs.

If many monosemous brother words are found in the fourth or third level, there are many candidate morphemes to choose from. A further selection is made based on calculation of sense similarity. More similar brother words are chosen. A threshold may be decided beforehand through experiment.

3 VW-Based Unsupervised WSD Method

VW is a solution to the semantic knowledge acquisition problem, and it doesn't limit the choice of stochastic learning methods. The mathematical modeling methods can all be applied to VW-based WSD methods. This section focuses on the application of the VW concept to WSD, and chooses mutual information and Bayesian method for the classifier construction.

3.1 A Sense Classifier Based on the Bayesian Model

Because the language model acquires knowledge from the VWs but not from the original ambiguous word, this method introduced here doesn't need human tagging of training corpus.

In the training stage for WSD, statistics of VWs and context words are obtained and stored in a database for next stage. Senseeval-3 data set plus unsupervised learning method are adopted to investigate into the value of VW in WSD. To ensure the comparability of experiment results, a Bayesian classifier is used in the experiments.

3.1.1 Bayesian Classifier

According to present research results, we know that the Bayesian classifier is an ideal WSD model, which is simple, efficient, and shows good performance. The classifier in this paper is as follows:

$$S(w_i) = \arg\max_{S_k} \left[\log P(S_k) + \sum_{v_j \in c_i} \log P(v_j \mid S_k) \right]$$

Where w_i is the ambiguous word, $P(S_k)$ is the occurrence probability of the sense S_k, $P(v_j|S_k)$ is the conditional probability of the context word v_j, and c_i is the collection of the context words.

To simplify the experiment process, the Naïve Bayesian modeling is adopted for the sense classifier. Feature selection and ensemble classification are not applied, which is both to simplify the calculation and to reveal the function of VWs in WSD.

3.1.2 Experiment Setup and Result

The Senseval-3 Chinese ambiguous words are taken as the test set, which includes 20 words, each with 2-8 senses. The data are divided into a training set and a test set by ratio of 2:1. There are 15-20 training instances for each sense of the words, and occurs by the same frequency in the training and test set.

Supervised WSD is first implemented using the Bayesian model on the Senseval-3 data set. With a context windows of (–10, + 10), the open test results are shown in Table 2.

The criterion items in Table 2 are defined as follows:

$$F = \frac{2P \times R}{P + R}$$

In which C(tagged) is the tag numbers, C(correct) is the number of correct tags, and C(all) is the number of all tagging in the gold standard set. Every sense of the ambiguous word has a P value, an R value and an F value. The F-value in table 2 is a weighted average of all the senses.

In the VW-based unsupervised WSD experiment, a 100M corpus (People's Daily for year 1998) is used for the VW training instances. The Senseval-3 data is used for the test. The experiment results with a context window of (–10, +10) and no feature selection is as in Table 3.

Table 2. The results in the supervised WSD

Ambiguous word	citation (Qin and Wang, 2005)	F-measure	Ambiguous word	citation (Qin and Wang, 2005)	F-measure
把握(bao3 wo4)	0.56	0.87	没有(mei2 you3)	0.75	0.68
包(bao1)	0.59	0.75	起来(qi3 lai2)	0.82	0.54
材料(cai2 liao4)	0.67	0.79	钱(qian2)	0.75	0.62
冲击(chong1 ji2)	0.62	0.69	日子(ri4 zi3)	0.75	0.68
穿(chuan1)	0.80	0.57	少(shao3)	0.69	0.56
地方(di4 fang1)	0.65	0.65	突出(tu1 chu1)	0.82	0.86
分子(fen1 zi3)	0.91	0.81	研究(yan2 jiu1)	0.69	0.63
运动(yun4 dong4)	0.61	0.82	活动(huo2 dong4)	0.79	0.88
老(lao3)	0.59	0.50	走(zou3)	0.72	0.60
路(lu4)	0.74	0.64	坐(zuo4)	0.90	0.73
Average	0.72	0.69	Note: Average of the 20 words		

Table 3. The results in unsupervised WSD based on VWs

sequence Number	Ambiguous word	F-measure	sequence Number	Ambiguous word	F-measure (%)
1	把握(bao3 wo4)	0.93	11	没有(mei2 you3)	1.00
2	包(bao1)	0.74	12	起来(qi3 lai2)	0.59
3	料(cai2 liao4)	0.80	13	钱(qian2)	0.71
4	冲击(chong1 ji2)	0.85	14	日子(ri4 zi3)	0.62
5	穿(chuan1)	0.79	15	少(shao3)	0.82
6	地方(di4 fang1)	0.78	16	突出(tu1 chu1)	0.93
7	分子(fen1 zi3)	0.94	17	研究(yan2 jiu1)	0.71
8	运动(yun4 dong4)	0.94	18	活动(huo2 dong4)	0.89
9	老(lao3)	0.85	19	走(zou3)	0.68
10	路(lu4)	0.81	20	坐(zuo4)	0.67
Average	0.80	Note: Average of the 20 words			

3.2 Experiment Analysis and Discussion

3.2.1 Experiment Evaluation Method

Two evaluation criteria are used in the experiments, which are the F-measure and precision in table 2 and table 3. Precision is a usual criterion in WSD performance analysis. Only in recent years the precision, recall and F-measure are taken to evaluate the WSD performance.

A comparison of the precision and F-measure is made in this paper. From table 2 and table 3, we can see that they do not vary much from each other. In supervised WSD, the difference between the two criteria is smaller than 2 percent, and in unsupervised WSD, only a 0.5 percent difference exists. On the whole, they have a similar value and tendency, which verifies that it's not amiss to use precision only in WSD performance analysis.

3.2.2 Result Analysis on Bayesian Supervised WSD Experiment

The experiment reveals that the results of supervised WSD and those of reference [11] are different. Though they are all based on the Bayesian model, the reference has used an ensemble classifier and there may be implementation differences. From the average value, the difference is not remarkable.

A supervised reference system is devised for the unsupervised WSD experiment, as in table 2. Gale suggests a lower bound or baseline value for WSD performance [12] Lower bound refers to the WSD precision of the simplest method, usually the percentage of the maximum frequency. Without the lower bound, we cannot decide the real difficulty of the WSD task. For example, for a two-sense ambiguous word, a 90% precision is quite remarkable if the frequency distribution is 1:1 between the two senses, but is quite humble for a sense frequency distribution of 9: 1. So we have to look at the precision in a scientific view.

As introduced above, in the supervised WSD experiment, the various senses of the instances are evenly distributed. The lower bound as Gale suggested should be very low and it's more difficult to disambiguate if there are more senses. The experiment verifies this reasoning, because the highest F-measure is less than 90%, and the lowest is less than 60%, averaging about 70%.

With the same number of senses and the same scale of training data, there is a big difference between the WSD results. This shows that other factors exist which influence the performance other than the number of senses and training data size. For example, the discriminability among the senses is an important factor. The WSD task becomes more difficult if the senses of the ambiguous word are more similar.

3.3 Experiment Analysis of the VW-Based WSD

The VW-based unsupervised method has taken the same open test set as the supervised method. The unsupervised method shows a better performance, with the highest precision at 100%, lowest over 60% and average over 80%. The results ensure that VW is useful in unsupervised WSD method.

But the average value of the 20 words is remarkably lower than in the reference. The reason why there's such a big difference between the results of the same method lies partly in the different test set, partly in details of the system implementation. On

the other hand, the former WSD experiment has adopted the natural distribution of senses in the corpus, which may lead to unbalanced distribution and lower difficulty of the WSD task. The experiment of the table 3 is under an even distribution of senses, which, as discussed above, will lead to a higher difficulty of the WSD task.

Compare the results in table 2 and table 3, we find that 16 among the 20 ambiguous words show better WSD performance in unsupervised than in supervised WSD, while only 2 of them shows similar and 2 worse performance. The average performance of the unsupervised method is higher by more than 10%. The reason lies in the following aspects:

1) Because there are several morpheme words for every sense of the word in construction of the VW, rich semantic information can be acquired in the training step and is an advantage for sense discrimination.

2) Senseval-3 has provided a small-scale training set, with 15-20 training instances for each sense, which is not enough for the WSD modeling. The lack of training information leads to a low performance of the supervised methods.

3) With a large-scale training corpus, the unsupervised WSD method has got plenty of training instances for a high performance in discrimination.

4) The discriminability of some ambiguous word may be low, but the corresponding VWs could be easier to disambiguate. For example, the ambiguous word “穿” has two senses which are difficult to distinguish from each other, but its VW’s senses of “越过/穿过/穿越” and “戳/捅/通/扎” much easier. This is also the case for the word “冲击”, whose VW’s senses are “撞击/磕碰/碰撞” and “损害/伤害”. VW-based knowledge acquisition for WSD of these ambiguous words has helped a lot to a high performance.

4 Conclusion

As discussed above, the supervised WSD method shows a low performance because of its dependency on the size of the training data. This reveals its weakness in knowledge acquisition bottleneck. VW-based unsupervised method has overcome this weakness and needs no manual corpus tagging for a satisfactory WSD performance. VW-based method is a promising solution to large-scale WSD task.

Acknowledgement. The authors would like to thank the anonymous reviewers for their helpful comments. This work is supported by the National Natural Science Foundation of P.R. China (Grant: 60603092) and the Ph.D. Programs Foundation of Ministry of Education of China (Grant: 20070217043).

References

1. Roberto, N.: Meaningful Clustering of Senses Helps Boost Word Sense Disambiguation Performance. In: Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL, Sydney, July 2006, pp. 105–112 (2006)
2. Peter, F.B., Stephen, A.: Della, P.: Word-Sense Disambiguation Using Statistical Methods. In: Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics, pp. 264–270 (1991)

3. Lyle, U., Martha, P.: An Empirical Study of the Behavior of Active Learning for Word Sense Disambiguation. In: Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL, New York, June, pp. 120–127 (2006)
4. Zhu, J.B., Eduard, H.: Active Learning for Word Sense Disambiguation with Methods for Addressing the Class Imbalance Problem. In: Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Prague, June, pp. 783–790 (2007)
5. Dong, Z.Z.: (2007), http://www.keenage.com
6. Li, C., Li, H.: Word Translation Disambiguation Using Bilingual Bootstrapping. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, pp. 343–351 (2002)
7. William, A.G., Church, K.W., Yarowsky, D.: Word on Statictical Methods for Word Sense Disambiguation. In: Goldman, R., Norvig, P., Charniak, E., Gale, B. (eds.) Working Notes of the AAAI Fall Symposium on Probabilistic Approaches to Natural Language, pp. 54–60. AAAI Press, Menlo Park (1992)
8. Tanja, G.: Statistical Corpus-Based Word Sense Disambiguation: Pseudowords vs. Real Ambiguous Words. In Companion Volume to the. In: Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL/EACL 2001) – Proceedings of the Student Research Workshop, Toulo (2001)
9. Preslav, N., Marti, H.: Category-based Pseudowords. In: the Companion Volume of the Proceedings of HLT-NAACL 2003, Edmonton, Canada (May 2003)
10. Henrich, S.: Automatic Word Sense Discrimination. Computational Linguistics 24(1), 97–123 (1998)
11. Qin, Y., Wang, X.J.: A Track-based Method on Chinese WSD. In: Joint Symposium of Computational Linguistics of China (JSCL 2005), China, NanJIng (2005)
12. William, A.G., Church, K.W., Yarowsky, D.: Estimating Upper and Lower Bounds on the Performance of Word Sense Disambiguation Programs. In: Proceedings of the 80th Annual Meeting of the Association for Computational Linguistics, Newark, Delaware (1992)

New Algorithm for Determining Object Attitude Based on Trapezoid Feature*

Lijuan Qin, Yulan Hu, Yingzi Wei, Hong Wang, and Yue Zhou

School of Information Science and Engineering, Shenyang Ligong University, Shenyang, Liaoning Province, China
qinlijuan1978@gmail.com

Abstract. Trapezoid is one feature which is widely presented in natural environments. In this paper, we propose a new pose estimation method based on trapezoid. The algorithm is simple to solve. It has good real-time characteristic. Furthermore, there exists unique solution for object pose from rectangle to locate. Thus, it has good practical value in application. At last, simulation results show this method is reasonable and has good real-time characteristic.

Keywords: Attitude; Trapezoid; Closed-form solution; Real-time characteristic.

1 Introduction

Attitude determination of an object with respect to a camera is an important problem in computer vision. It has wide varieties of applications, such as space mobile robots, robot self-positioning, camera calibration, hand-eye calibration and pose identification of known 3D object.

Points, lines and high-level features (circles, ellipses and conic sections, etc.) are the most popular features for pose estimation. For pose estimation from point correspondences, many researchers have done much work on it. The representative papers include paper [1~3]. Line features is easier to extract and the accuracy of line detection is higher than point because the images we obtain are often unclear. Thus pose estimation from line correspondences is widely discussed by many researchers. Pose estimation method from line features has made great progress on it. Literature [4~5] presented the iterative method to solve. Literature [6~7] introduced closed-form method to solve. High-level features include circles, ellipses and conic sections, etc. The representative papers for pose estimation from high-lever features include [8~9].

We usually see many pose estimation methods from point features, line features and high-lever features. However, we have not seen much research on pose estimation from trapezoid feature. Thus, research on attitude determination from trapezoid feature has some theoretical value. In the view of practice, trapezoid is one feature which is widely presented in natural environments. Therefore, using trapezoid feature to locate has important practical value.

* Supported by National High Technology Research and Development Program of China under grant No. 2003AA411340.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 405–413, 2008.
© Springer-Verlag Berlin Heidelberg 2008

We device a new method to determine object pose from trapezoid feature. This algorithm is simple to solve. It has good real-time character. Therefore it can be included in a runtime visual process. In addition, it can obtain unique solution. At last, simulation results demonstrate this method is sound and has good real-time characteristic.

The remainder of this paper is organized as follows: Section 2 presents a formal statement of attitude determination from trapezoid feature. In section 3, new closed-form approach based on trapezoid feature is presented. Section 4 shows the rigid transformation result from the object frame to the camera frame. In Section 5, we present our experimental results. Section 6 is the conclusion.

2 Problem Formulation

Attitude determination based on trapezoid feature can be described as follows: we know the positional information of trapezoid feature in the object frame and we know the corresponding trapezoid feature in the image plane. The problem is to determine the rigid transformation between the object frame and the camera frame (see Fig.1).

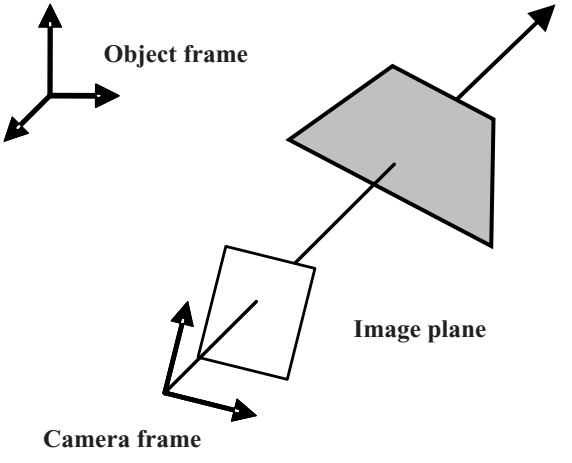


Fig. 1. General formulation of attitude determination

This problem can be decomposed in two parts: At first, from image trapezoid in image plane and the geometric information of trapezoid feature, we want to know the pose of trapezoid in the camera frame. In a second step, we compute the rigid transformation between the object frame and the camera frame.

3 Object Attitude from Trapezoid Feature

We assume the parameters of the projection (the intrinsic camera parameters) are known and we consider a pin-hole camera model. As shown in Fig. 2, we assume that vectors of space lines $L_i\,(i=1\sim 4)$ on trapezoid in the object frame are $V_i=(A_{wi},B_{wi},C_{wi})$.

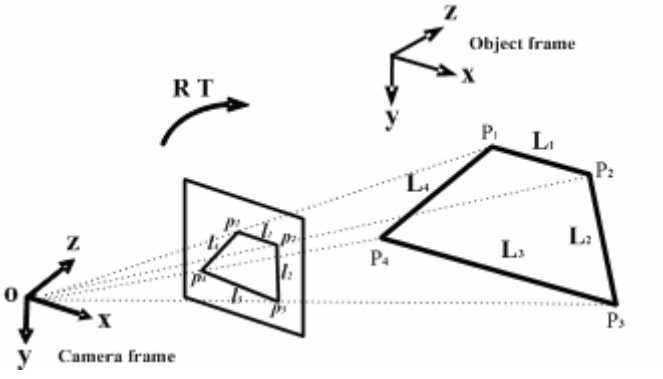


Fig. 2. Perspective projection of trapezoid

We assume four space lines intersect at four points. And the coordinates of four points are (X_{wi}, Y_{wi}, Z_{wi}) in the object frame. The vectors of four lines in the camera frame are $V_i = (A_i, B_i, C_i)$. The coordinates of points P_i in the camera frame are (X_i, Y_i, Z_i). We know the image line equation is as follows: $a_i x + b_i y + c_i = 0$ and we consider an image line l_i characterized by a vector $v_i(-b_i, a_i, 0)$ and a point t_i of coordinates (x_i, y_i, f). The coordinates of image points p_i for four points P_i are (x_i, y_i, f). In addition, we assume the direction vector of line L_2 is from point p_2 to point p_3 and the direction vector of line L_4 is from point p_1 to point p_4. The problem to determine object attitude is to determine the rotation and translation matrix between the object frame and the camera frame.

At the same time, we also know geometrical constraints as follows:

(a) Line L_1 is parallel to line L_3. (b) The angle between line L_1 and line L_2 is α and the angle between line L_1 and line L_4 is β. (c) The angle between line L_2 and x axis is an acute angle and the angle between line L_4 and x axis is an obtuse angle. (d) The length of line L_1 is $|P_1P_2| = a$. (e) The distance between point P_2 and the optical center is bigger than the distance between point P_1 and the optical center.

As shown in Fig. 2, the perspective projection model constraints line L_i, image line l_i and the origin of the camera frame to lie in the same plane. This plane is called projection plane S_i. The vector N_i normal to this plane can be computed as the cross product of ot_i and vector v_i. Thus, we get:

Then
$$N_i = ot_i \times v_i = \begin{pmatrix} x_i \\ y_i \\ f \end{pmatrix} \times \begin{pmatrix} -b_i \\ a_i \\ 0 \end{pmatrix} = \begin{pmatrix} a_i f \\ b_i f \\ c_i \end{pmatrix} \quad (1)$$

From constraints (a) that line L_1 is parallel to line L_3, and line L_3 is perpendicular to normal vector N_3, thus line L_1 is perpendicular to normal vector N_3. We know line L_1 is perpendicular to normal vector N_1, therefore the vector of line L_1 can be computed by the cross product of the vector of N_1 and N_3, we have:

$$\begin{pmatrix} A_1 \\ B_1 \\ C_1 \end{pmatrix} = N_1 \times N_3 = \begin{pmatrix} a_1 f \\ b_1 f \\ c_1 \end{pmatrix} \times \begin{pmatrix} a_3 f \\ b_3 f \\ c_3 \end{pmatrix} = \begin{pmatrix} b_1 f c_3 - b_3 f c_1 \\ -a_1 f c_3 + a_3 f c_1 \\ a_1 b_3 f^2 - a_3 b_1 f^2 \end{pmatrix} \tag{2}$$

For line L_2, because the norm of the unit direction vectors of line L_2 is unit and line L_2 lies in the interpretation plane formed by the image line l_2 and the optical centre, at the same time, the angle between line L_3 and line L_2 is α from constraints (b), we can obtain the following three equations:

$$\begin{aligned} &A_2 N_{21} + B_2 N_{22} + C_2 N_{23} = 0 \\ &A_1 A_2 + B_1 B_2 + C_1 C_2 = \cos(\alpha) \\ &A_2^{\,2} + B_2^{\,2} + C_2^{\,2} = 1 \end{aligned} \tag{3}$$

Thus, A_2, B_2, C_2 can be expressed by A_1, B_1, C_1.

From constraint (c) that the angle between line L_2 and x axis is an acute angle, we can draw the conclusion that the value of A_2 is positive, thus we have:

$$\begin{aligned} A_2 &= \frac{-(mn + pq) + \sqrt{(mn + pq)^2 - (n^2 + q^2 + 1)(p^2 + m^2 - 1)}}{(n^2 + q^2 + 1)} \\ B_2 &= m + nA_2 \\ C_2 &= p + qA_2 \end{aligned} \tag{4}$$

From equation (4), we obtain the unique solution of the direction vectors of line L_2. For line L_4, because the norm of the unit direction vectors of line L_4 is unit and line L_4 lies in the interpretation plane formed by the image line l_4 and the optical centre, at the same time, the angle between line L_4 and line L_1 is β from constraints (b), we can obtain the following three equations:

$$\begin{aligned} &A_4 N_{41} + B_4 N_{42} + C_4 N_{43} = 0 \\ &A_4 A_1 + B_4 B_1 + C_4 C_1 = \cos(\beta) \\ &A_4^{\,2} + B_4^{\,2} + C_4^{\,2} = 1 \end{aligned} \tag{5}$$

In above equations, A_4, B_4, C_4 can also be expressed by A_1, B_1, C_1.

From constraint (c) that the angle between line L_4 and x axis is an obtuse angle, we can draw the conclusion that the value of A_4 is positive, thus we have:

$$\begin{aligned} A_4 &= \frac{-(gh+wl)-\sqrt{(gh+wl)^2-(h^2+l^2+1)(w^2+g^2-1)}}{(h^2+l^2+1)} \\ B_4 &= g+hA_3 \\ C_4 &= w+lA_3 \end{aligned} \tag{6}$$

From equation (6), we obtain the unique solution of the direction vectors of line L_4. For line L_1, the vector of line L_1 can be defined by point P_1 and point P_2 on it, we have:

$$L_1 = \begin{pmatrix} A_1 \\ B_1 \\ C_1 \end{pmatrix} = \begin{pmatrix} k_2x_2 - k_1x_1 \\ k_2y_2 - k_1y_1 \\ k_2z_2 - k_1z_1 \end{pmatrix} \tag{7}$$

Equation (2) and (7) is equal, we have:

$$A_2(k_2x_2 - k_1x_1) + B_2(k_2y_2 - k_1y_1) + C_2(k_2z_2 - k_1z_1) = a\sqrt{A_2^2 + B_2^2 + C_2^2} \tag{8}$$

From (8), we get the expression about variable k_1 and k_2:

$$(A_2x_2 + B_2y_2 + C_2z_2)k_2 - (A_2x_1 + B_2y_1 + C_2z_1)k_1 = a\sqrt{A_2^2 + B_2^2 + C_2^2} \tag{9}$$

We assume:

$$k_2 = s_1k_1 + s_2 \tag{10}$$

where, $$s_1 = \frac{A_2x_1 + B_2y_1 + C_2z_1}{A_2x_2 + B_2y_2 + C_2z_2} \quad s_2 = \frac{a\sqrt{A_2^2 + B_2^2 + C_2^2}}{A_2x_2 + B_2y_2 + C_2z_2}$$

From constraint (d) the length of line L_1 is $|P_1P_2| = a$, we get:

$$\begin{aligned} &\left(\frac{k_2x_2}{\sqrt{(x_2^2 + y_2^2 + z_2^2)}} - \frac{k_1x_1}{\sqrt{(x_{c1}^2 + y_{c1}^2 + z_{c1}^2)}}\right)^2 + \left(\frac{k_2y_2}{\sqrt{(x_2^2 + y_2^2 + z_2^2)}} - \right. \\ &\left. \frac{k_1y_1}{\sqrt{(x_1^2 + y_1^2 + z_1^2)}}\right)^2 + \left(\frac{k_2z_{c2}}{\sqrt{(x_2^2 + y_2^2 + z_2^2)}} - \frac{k_1z_{c1}}{\sqrt{(x_1^2 + y_1^2 + z_1^2)}}\right)^2 = a^2 \end{aligned} \tag{11}$$

which can be expressed as follows:

$$f_1k_2^2 + f_2k_1k_2 + f_3k_1^2 = 0 \tag{12}$$

Substituting (10) into (12), we get the expression about variable k_1:

$$f_4k_1^2 + f_5k_1 + f_6 = 0 \tag{13}$$

From (13), we get:

$$k_1 = \frac{-f_5 \pm \sqrt{f_5^2 - 4 f_4 f_6}}{2 f_4} \tag{14}$$

When

$$k_1 = \frac{-f_5 + \sqrt{f_5^2 - 4 f_4 f_6}}{2 f_4} \tag{15}$$

Then substituting (15) into (10), we have:

$$k_2 = s_1 + s_2 \frac{-f_5 + \sqrt{f_5^2 - 4 f_4 f_6}}{2 f_4} \tag{16}$$

When

$$k_1 = \frac{-f_5 - \sqrt{f_5^2 - 4 f_4 f_6}}{2 f_4} \tag{17}$$

Then substituting (17) into (10), we obtain:

$$k_2 = s_1 + s_2 \frac{-f_5 - \sqrt{f_5^2 - 4 f_4 f_6}}{2 f_4} \tag{18}$$

Therefore, we get two groups of closed-form solutions about two variables of k_1 and k_2, whose physical meaning is the distance between points P_i $i = (1 \sim 2)$ and the optical centre. There is one way to get the unique solution of k_1 and k_2. We can control the placement position of Trapezoid to eliminate wrong solutions. In this method, we make the distance between point P_2 and the optical center is bigger than the distance between point P_1 and the optical center. That is constraint (e). From constraint (e), we compare (15) and (16), (17) and (18). Then, we can get one solution about two variables of k_1 and k_2.

Thus, the coordinates of point P_1 $i = (1 \sim 2)$ in the camera frame can be determined. In addition, the vectors of L_i $(i = 1, 2, 4)$ have been determined from equation (2), (4) and (6). Thus the pose of Trapezoid with respect to the camera can be located uniquely. Then, the rigid transformation between the object frame and the camera frame can be determined using rotation theory.

4 Compute Rigid Transformation

Let us image that the camera frame is obtained by first rotating the object frame with rotation matrix R then translating it with vector T. Let $V_i' = (Aw_i', Bw_i', Cw_i')$ be the vectors of line L_i in the object frame and its camera vector representations is:

$$V_i = RV_i' \tag{19}$$

Rotation matrix R can be expressed by the tilt, swing and spin angles. If we consider the nine elements of R as independent unknowns, then equation (19) is a homogeneous linear equation. Therefore, with three 2D to 3D line correspondences, these nine unknowns can be obtained by solving linear equations.

From computation in section 3, we have obtained the vectors of lines $L_i\ (i=1,2,4)$ in the camera frame. Thus, with three 2D to 3D line correspondences, matrix R can be computed easily from equation (19).

A translation vector can be expressed as follows:

$$T = \begin{pmatrix} T_x & T_y & T_z \end{pmatrix}'$$

$$P = RP' + T \tag{20}$$

R and T can be determined separately. Therefore, the translation vector can be uniquely determined by one point P_1 correspondence on trapezoid.

5 Experimental Results

In our simulation experiments, we choose the concrete parameters of the camera are $\begin{pmatrix} 800 & 0 & 0 \\ 0 & 800 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The size of image is 512×512. The view angle of camera is about 36×36 degrees. The trapezoid is chosen such that its image is within the view of the camera.

To test the legitimacy of algorithm for pose estimation from trapezoid, we do some simulation experiments. The results for pose estimation based on trapezoid feature are as table 1 shown.

From experimental results, we can see the closed-form method based on trapezoid we have introduced in this paper works well and can be used to locate object.

Table 1. Pose estimation results based on trapezoid feature

Case	image line l_1	image line l_2	image line l_3	image line l_4	α	β	γ	Tx	Ty	Tz
1	a_1=0.0338	a_2=-0.9821	a_3=-0.0303	a_4= 0.9787	1	2	-2	1	4	300
	b_1= 0.9804	b_2= 0.0348	b_3=-0.9788	b_4=-0.0364						
	c_1=-0.1298	c_2= 0.0277	c_3= 1.0919	c_4= 0.9333						
2	a_1=-0.0725	a_2=-1.4006	a_3= 0.0459	a_4= 0.9481	3	-2	3	2	6	310
	b_1= 1.4107	b_2=-0.0745	b_3=-0.9496	b_4= 0.0470						
	c_1=-0.3902	c_2= 0.1523	c_3= 1.0775	c_4= 0.8378						
3	a_1=-0.0156	a_2=-0.9192	a_3= 0.0110	a_4= 0.9237	2	-3	1	3	4	320
	b_1= 0.9260	b_2=-0.0180	b_3=-0.9230	b_4= 0.0151						
	c_1=-0.1126	c_2= 0.0871	c_3= 0.9613	c_4= 0.7685						
4	a_1=-0.0153	a_2=-0.8994	a_3= 0.0139	a_4= 0.9008	-3	-1	1	1	8	330
	b_1= 0.8964	b_2=-0.0147	b_3=-0.9007	b_4= 0.0190						
	c_1=-0.2136	c_2= 0.0304	c_3= 1.0250	c_4= 0.7763						

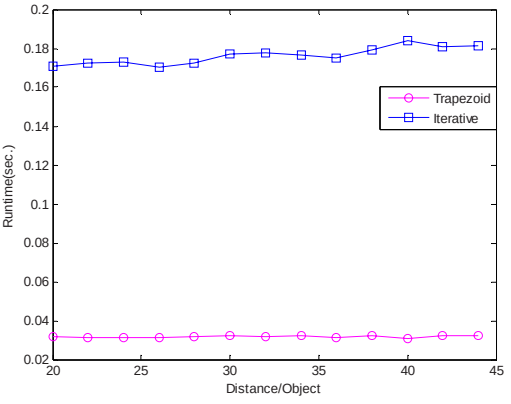


Fig. 3. Comparison result of runtime

For testing the real-time characteristic of the algorithm we have introduced, we do the contrastive experiments with iterative method in literature [7]. In Fig. 3, the horizontal coordinates are degrees. From experimental results (as shown in Fig. 3), we find the runtime of the closed-form method based on trapezoid is shorter than that of the iterative method [7], which shows the closed-form method based on trapezoid feature we have proposed has good real-time characteristic.

6 Conclusion

Determining the attitude of an object with respect to a camera is an important problem in computer vision. Trapezoid is one feature which is widely presented in man-made environments. In this paper, we propose a new pose estimation method based on trapezoid feature. This method is simple and convenient. It has good real-time characteristic. Furthermore, there exists unique solution for object pose from trapezoid to locate. Thus, this algorithm has good practical value in application. It can be applied to robot self-positioning, pose identification of known 3D object and camera calibration.

References

1. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6), 381–395 (1981)
2. Wolfe, W.J., Mathis, D., Sklair, C.W., Magee, M.: The perspective view of three points. IEEE Transactions on Pattern Analysis and Machine Intelligence 13(1), 66–73 (1991)
3. Haralick, R.M., Lee, C., Ottenberg, K., Nolle, M.: Analysis and solutions of the three point perspective pose estimation problem. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 592–598 (1991)
4. Yuan, J.S.-C.: A general photogrammetric method for determinating object position and orientation. IEEE Trans.Pattern Anal.Machine Intell. 5(2), 129–142 (1989)

5. Lowe, D.: Fitting parameterized three-dimensional models to images. IEEE Trans.Pattern Anal.Machine Intell. 13(5), 441–450 (1991)
6. Dhome, M., Richetic, M., Lapreste, J.-T., Rives, G.: Determination of the attitude of 3-D objects from a single perspective view. IEEE Trans.Pattern Anal.Machine Intell. 11(12), 1266–1278 (1989)
7. Chen, H.: Pose determination from line to plane correspondences: existence solution and closed form solutions. IEEE Trans.Pattern Anal.Machine Intell. 13(6), 530–541 (1999)
8. Shiu, Y.C., Zhuang, H.: Pose determination of cylinders, cones and spheres from perspective projections. In: Applications of Artificial Intelligence X: Machine Vision and Robotics, pp. 771–782 (1992)
9. Haralick, R.M.: Solving camera parameters from perspective projection of a parameterized curve. Pattern Recognition 17(6), 637–645 (1984)

Curvature Feature Based Shape Analysis

Yufeng Chen, Fengxia Li, and Tianyu Huang

Beijing Laboratory of Intelligent Information Technology
School of Computer Science and Technology, Beijing Institute of Technology*
100081 Beijing, P.R. China
yfchen@bit.edu.cn

Abstract. A novel method is presented to evaluate the similarity of shapes based on the curvature features and their distribution. Firstly the curvature information is used to define the curvature features, which are learned and searched by the proposed statistic method. Secondly the structural features of each pair are measured, so that the distribution of the curvature feature can be further measured. Taking both advantages of the local shape context analysis and the global feature distances optimization, our method can endure large nonrigid distortion and occlusion. The experiments, which have been implemented on the MPEG-7 shape database, show that this method is efficient and robust under certain shape distortion. Another experiment on the abnormal behavior detection shows its potential in shape detection, motion tracking, image retrieving and the related areas.

1 Introduction

Shape representation and matching is important in computer vision and image processing, since it contains the main information of the object. The solutions are significant with the increasing requirement of content-based image retrieval, detection, recognition, motion tracking and other applications.

Many methods have been reported to solve the problems, of which there are mainly two types: region-based methods such as shape moment, and contour-based methods such as chain code, More details and comparisons of these methods are introduced by Veltkamp[1] and Zhang[2].

The most related method to ours is the Curvature Scale Space(CSS) method[3]. Firstly they turn the curvature point of different scales into a scale space map, and then the shapes are compared by the difference of point trajectory. CSS method is characterized by invariance to scaling and rotation, and efficient to represent, however, the main problem lies in that it fails in captureing the global features or the shape details, at the same time the matching is too complicated[4].

Many other analysis methods have been performed on the contours. The most basic ones are frequency analysis[5] and wavelet method[6]. As Zhang[2] has

* This project was supported by the National Defense Basic Scientific Research Grant No. B2220061084 and China Postdoctoral Science Foundation No. 20070410464.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 414–421, 2008.
© Springer-Verlag Berlin Heidelberg 2008

pointed out the frequency result may be affected by local distortion, while the wavelet method is too complicated to get better representation in an efficient way.

Shape context[7] methods such as Multi-scale Convexity Concavity(MCC)[8] are other ways to take into account of more details. However, most of them are considered globally, which can endure limited distortion, and the optimizations are often locally optimized[9]. The contradiction of representation and and matching between the global and the local still exists.

Although many methods[10] have been developed to consider the similarity of the context and its distribution, the result is highly depends on the similarity and integrity of the contour.

Recently, more and more efforts are made on the local features to analyze shape. Some results[11] show that local features can represent shape efficiently and accurately. And some local shape detection methods[12]are proposed, which try to detect the existing curve and are compared with models to understand shapes. However, how to define the local shape feature and analyze their distribution is not answered yet, which is very important for the shape structure analysis.

Aiming to solve the problem, we propose a feature-based shape matching method. Our method employs curvature analysis and geometry similarity to consider both local shape contexts and global matching.

2 Curvature Based Feature Segmentation

2.1 Differential Expression of Curve

For any differentiable regular curve in the R^2 Euclidean space, we can derive a transformation from $r{:}[a,b] \in R$ to R^2, where, r is parameter curve. Denote the tangent vector: $\dot{r(t)}$, then the length of the curve could be expressed as:

$$L(a,b) = \int_a^b |\dot{r(t)}|dt,$$

which is independent of rigid transformation, thus $x = |\dot{r(t)}|$ can be the better choice of curve parameter.

To establish the local coordinates, we denote

$$\alpha(x) = \dot{r(x)}, \qquad \beta(x) = \dot{\alpha(x)},$$

which are the Frenet Frame in R^2, and stand for the tangent vector and the curvature vector, respectively.

According to the properties of the derivative, the differential expressions represent a family of curves with different border conditions. So $\beta(x)$ is the ideal form for the curves, which is intrinsically invariant to transfer and rotation. Given the coordinate and the direction it can represent any curve of the family.

Two advantages can be obtained from the new expression of the curve. Firstly, it is invariant under transfer and rotation; secondly it can handle the non-rigid shape change and be used for local analysis.

2.2 Curve Segmentation

From point of view of human vision, the shapes differ from the distribution of the border points which is greatly affected by the scale. The scale space[13] is introduced to give a full view of the shape, based on which, the Curvature Scale Space(CSS) method are successfully used in MPEG-7 for shape description[3,14].

The scale factor s, which stands for the size of the filter windows, is used to define the filter $flt(\cdot)$ of different scales. The family of filtered contour under different scales is shown in Fig.1.(a),(b) and (c).

$$\beta(x, s) = flt(\beta(x), s).$$

Feature points are always needed in the local analysis of the shape, since most of the work is done directly on the curvature of the shape such as inflexion point used in CSS and extreme points used in structure analysis, see Fig.1.(d)and(e).

The main limitation of the curvature point is that it can not describe the convex part of the shape as shown in Fig.1.(f). To the extreme, circle eclipse and all kinds of convex polygons have no such feature points therefore they can be considered as the same by mistake.

To avoid these problems, the curve energy function of the shape is introduced for segmentation.

$$En(x, s) = H(\beta(x, s)).$$

Our feature points are defined on these functions.

$$\begin{aligned} X_b(s) = \{x_{bj} : & \frac{d}{dx} En(x_{bj}, s) = 0, \\ & \frac{d^2}{dx^2} En(x_{bj}, s) >= 0, x_{bi} < x_{bj}, if, i < j\}. \end{aligned}$$

where x_{bj} is the minimum point, $i, j < N_b, i, j \in N$ and N_b is the number of minimum points, then the segment are specified.

$$\begin{aligned} W(a, b) &= \begin{cases} 1, x \in [a, b] \\ 0, \quad else \end{cases} \\ Seg(s, a, d) &= En(s) \times W(a, a + d) \\ FTR(s) &= \{Seg(s, a_j, d_j) : a_j = x_{bi}, \\ d_j &= x_{b(i+1)} - x_{bi} \int Seg(s, a_j, d_j)dx \geq \int Seg(s, a_m, d_m)dx, if, j < m\} \\ IR(s, j) &= \frac{\int Seg(s, a_j, d_j)dx}{\int Seg(s, 0, L)dx} \end{aligned}$$

where, $i, j, m < N_b \ i, j, m \in N \ x_{bi} \in X_b(s)$.

Note that the segments are ordered by their information rate IR. It is possible to record the main shape with N_n salient feature segments, which will accelerate the searching and matching process.

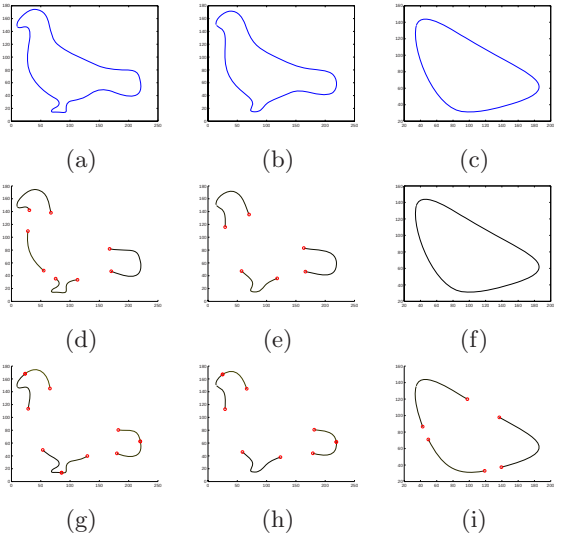


Fig. 1. Shape segmentation comparison. The contour of first column is under the filter size of 5% image width and the middle column is 10%, the last is 45%; the first row are the contour shapes under different scales, the second row are the main segmentation parts by curvature feature, the last are by our method.

It can be seen clearly from the Fig.1.(g),(h) and (i),that more segments are used and more details can be distinguished. Especially the convex shape can also be segmented into more reasonable shape features rather than structure features. This will be helpful in shape matching, since the structure may vary largely while the feature remains almost the same.

3 Curvature Feature Analysis

Once the feature segments are selected, the distribution of their curvature along the contour needs to be taken into account. Some characteristics of the segments should be considered and some requirements are needed to be met.

Being filtered under certain scale and segmented with curvature information, the segments are expected to be the simple curves with low frequency, therefore some simple methods such as statistics on the curvature are sufficient to evaluate the similarity of the segments.

To taking the advantage of geometry invariance, normalized central moments are preferred.

$$M_0(s,a,d) = \int_a^{a+d} En(s)dx \quad , \quad M = \frac{\int_a^{a+d} En(s)xdx}{M_0}$$

$$M_1(s,a,d) = \frac{1}{L(a,a+d)}M$$

$$M_k(s,a,d) = \frac{1}{L(a,a+d)^k}\int_a^{a+d} En(s)(x-M)^k dx, \quad x \in N, k \neq 1.$$

The first five orders are used here to stand for the main index of the feature, center, variance, skewness and so on.

The feature similarity distance is defined on the moment.

$$Fds(s,a,d,n) = \frac{1}{k}\sum_{i=0}^{k} \frac{M_i(s,a,d) - M_i'(s,a_n,d_n)}{max(M_i(s,a,d), M_i'(s,a_n,d_n))},$$

where k is the maximum moment order, n is the index of the feature, M' is the moment on the target shape, and a_n, d_n are the segment parameter of n-th feature defined in the function 1.

It is uncertain that which segment is the feature although the similarities are numerically displayed, since there may be more than one similar feature or some false features. The first few minimum points are selected as follows:

If the n-th feature of the target shape is searched, given the feature length d_n, the start point of the segment x_{ni} are defined at some minimum feature distance point on the object shape.

$$Xn = \{x_{ni} : \frac{d}{dx}Fds(s,x_{ni},dn,n) = 0, \frac{d^2}{dx^2}Fds(s,x_{ni},dn,n) >= 0,$$
$$Fds(s,x_{ni},dn,n) < sh, i <= N_m, i \in N\},$$

where, N_m is the candidate number of n-th similar feature, and sh is the threshold that can be learned from the samples.

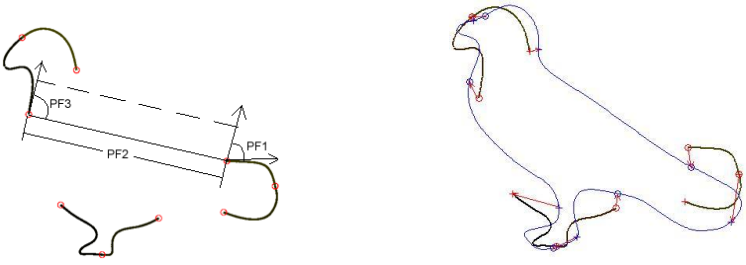
4 Shape Matching

To match the feature via their distribution is a typical optimization problem, many methods have been proposed. Dynamic programming[8,15] is used to register the point serial while geometric transformation[9] searched by EM algorithm is reported to match the points recently.

Different with these situations, our feature segments have more information to match than point set, such as curvature energy, direction, and their difference. Therefore, the searching complexity is largely reduce by the structure information between the features.

The structural features PF_i and the matching function $BD()$ between the pairs are defined (See Fig.2.(a)) to be invariant to transfer and rotation. The matching result of bird features in Fig.2.(a) is shown in Fig.2.(b).

$$PF_1(s,n,i,m,j) = \angle(r(x_{ni})) - \angle(r(x_{mj}))$$
$$PF_2(s,n,i,m,j) = ||r(x_{ni}) - r(x_{mj})||$$
$$PF_3(s,n,i,m,j) = \angle(r(x_{ni}) - r(x_{mj})) - \angle(r(x_{mj})).$$
$$BD(s,n,i,m,j) = \frac{1}{3}\sum_{i=0}^{3} \frac{PF_i(s,n,i,m,j) - PF_i'(s,n,1,m,1)}{max(PF_i(s,n,i,m,j) - PF_i'(s,n,1,m,1))}$$



(a) (b)

Fig. 2. Feature pair definition and matching

where $m, n <= N_n, m, n \in N, m \neq n$ are the feature index, and N_n is the number of salient features , $i, j \in N$ are the candidate ID of a certain feature, and x_{ni} is the start point of the feature segment.

From viewpoint of the statistic, the distance between features BD are changed with the distortion of the object shape. Then the matching probability of shape r can be deduced by minimizing the following function

$$F(r|BD) = \sum_{m,n,m\neq n} \{BD(s,n,i,m,j) \times \sqrt{Fds(s,x_{ni},d_n,n) \times Fds(s,x_{mj},d_m,m)}\}.$$

Therefore, the similarity of two shapes can be evaluated by searching maximum average value of these pair-matching-probability in the possible feature matching candidates.

5 Experiment

A serial of experiments have been carried out to show the properties of the novel method. Most of them are implemented on the MPEG-7 database to compare with some of the major methods such as FFT, CSS and MCC.

At first the comparison of the similarity evaluation between different shapes in Fig.3 is made to show the efficiency of feature searching and robustness in shape matching. Experiment shows that a throughout search with window size of 40 points, along bird shape contour of about 500 points by Short-time Frequency method takes about 3.1 seconds on the Matlab, which is as three times long as that with our method, and it will be longer for more complicated features as shown in Table.1.

Another experiment is carried out on 3 categories of the shape database, and about 10 samples in each category are selected for cross validate. The distances of samples are modeled first and the test image is then evaluated by the model. The result is compared with CSS and MCC reported by Adamek[8]. As we can see from Table 2, all of these methods are fit for the simple shapes, and the MCC

Table 1. Comparison of time consumed for feature searching along the contour under two different conditions

time(second)	SFFT Method	Our Method
500points	3.1	1.0
1000points	6.1	1.6

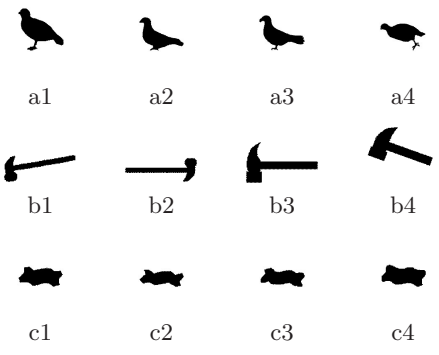


Fig. 3. Samples in MPEG-7 shape database

Table 2. Fig.3 Shape retrieval rate and comparison

%	bird	hammer	brick
CSS	75	65	100
MCC	75	92	100
OUR	83	90	100

method is more suitable for shapes with small shape changes, and our method is capable of dealing with small structure changes.

6 Conclusion

A novel shape matching method is proposed compared with some of the main methods reported in the previous literature. Experiments show that our method is suitable for the nonrigid matching under occlusion and incomplete matching. The method is prospected to be applied in surveillance, CBIR(Content Based Image Retrieval), detection, tracking and other related areas.

References

1. Veltkamp, R.C., Hagedoorn, M.: State-of-the-art in shape matching. Technichal Report UU-CS (Ext. r. no. 1999-27), Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands (1999)

2. Zhang, D., Lu, G.: Evaluation of MPEG-7 shape descriptors against other shape descriptors. Multimedia Systems 9(1) (2003)
3. Mokhtarian, F.: Silhouette-based isolated object recognition through curvature scale space. IEEE Transactions on Pattern Analysis and Machine Intelligence 17(5), 539–544 (1995)
4. Zhang, D., Lu, G.: A Comparison Of Shape Retrieval Using Fourier Descriptors And Short-Time Fourier Descriptors. In: Proceeding of The Second IEEE Pacific-Rim Conference on Multimedia, Beijing, China (October 2001)
5. Chellappa, R., Bagdazian, R.: Fourier Coding of Image Boundaries. IEEE Transactions on Pattern Analysis and Machine Intelligence 6(1), 102–105 (1984)
6. Tieng, Q., Boles, W.: Recognition of 2D Object Contours Using the Wavelet Transform Zero-Crossing Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(8), 910–916 (1997)
7. Thayananthan, A., Stenger, B., Torr, P.H.S., Cipolla, R.: Shape Context and Chamfer Matching in Cluttered Scenes. IEEE Conference on Computer Vision and Pattern Recognition, 127–133 (2003)
8. Adamek, T., Connor, N.: A Multiscale Representation Method for Nonrigid Shapes With a Single Closed Contour. IEEE Transaction on Circuits and Systems for Video Technology 14(5) (2004)
9. Tu, Z., Yuille, A.L.: Shape Matching and Recognition: Using Generative Models and Informative Features. In: European Conference on Computer Vision (2004)
10. Sebastian, T.B., Klein, P.N., Kimia, B.B.: On Aligning Curves. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(1), 116–124 (2003)
11. Stark, M., Schiele, B.: How Good are Local Features for Classes of Geometric Objects. In: Internatianl Conference of Computer Vision, Rio de Janeiro, Brazil (October 2007)
12. Opelt, A., Pinz, A., Zisserman, A.: A Boundary-Fragment-Model for Object Detection. In: European Conference on Computer Vision, Graz, Austria (May 2006)
13. Witkin, A.P.: Scale Space filtering. In: Proceeding of 8th International Joint Conference of Artificial Intelligence (1983)
14. Abbasi, S., Mokhtarian, F., Kittler, J.: Curvature scale space image in shape similarity retrieval. Multimedia Systems 7(6), 467–476 (1999)
15. Petrakis, E., Diplaros, A., Milios, E.: Matching and Retrieval of Distorted and Occluded Shapes Using Dynamic Programming. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(11) (2002)

A Study on the *P3P* Problem*

Jianliang Tang[1,2], Wensheng Chen[1] , and Jie Wang[3]

[1] College of Mathematics and Computational Science, Shenzhen University
Shenzhen, 518060, China
{jtang, chenws}@szu.edu.cn
[2] National Laboratory of Pattern Recognition, Institute of Automation
CAS, Beijing, 100080, China
[3] Key Laboratory of Multimedia and Intelligent Software
College of Computer Science and Technology, Beijing University of Technology
Beijing 100022, China

Abstract. In this paper, we focus on the number of solution for the Perspective-Three-point problem (*P3P*) in some geometrical constraints of 3 points, which is a common problem in applied mathematics and computer vision. We use Wu's zero decomposition method to find a complete triangular decomposition of a practical configuration for the *P3P* problem. By the Wu's method, we also obtain some sufficient conditions under which there are multi-solution for the *P3P* problem.

1 Introduction

The Perspective-*n*-Point (*PnP*) problem is an important and classic problem in applying mathematics and computer vision [1-3]. It is to determine the position and orientation of the camera with respect to a scene object from n correspondent points. It concerns many important fields, such as computer animation [8], robotics [1], etc. Fischler and Bolles [2] summarized the problem as follows:

``Given the relative spatial locations of *n* control points, and given the angle to every pair of control points from an additional point called the Center of Perspective (C_P), find the lengths of the line segments joining C_P to each of the control points."

The study of the *PnP* problem mainly consists of three subproblem: *P3P* problem, P4P problem and *P5P* problem because the problem is not well defined for $n<3$ and well solved by DLT method for $n>5$. The aim of this paper is to discuss the number of the solutions for the P3P problem.

The P3P problem is the most basic case of the *PnP* problems. It is also the smallest subset of control points that yields a finite number of solutions. All other *PnP* ($n>3$) problems include the *P3P* problem as a special case. The *P3P* problem has been well investigated. Fishler and Bolles [2] presented the RANSAC algorithm and showed that the *P3P* problem has at most four positive solutions and this upper bound is also attainable via a concrete example. DeMenthon et al [5, 6] showed that by using approximations to the perspective, simpler computational solutions could be obtained.

* Partially supported by the National Natural Science Foundation of China under Grant (No.10526031, 60603028)

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 422–429, 2008.
© Springer-Verlag Berlin Heidelberg 2008

One of the important research directions on the *P3P* problem is its multi-solution phenomenon. Fischler and Bolles [2] presented some examples of multi-solutions of the P3P problem. In 1991, Wolfe et al [4] gave a geometric explanation to this multi-solution phenomenon in the image plane under the assumption of ``canonical view.'' In 1997, Su et al [8] applied Wu's zero decomposition method to find the main solution branch and some non-degenerate branches for the *P3P* problem. But a complete decomposition was not given. In [7], they used the Sturm sequence to give some conditions to adjudicate the number of solutions. In 1998, Yang [9] gave partial solution classifications of the *P3P* problem under some non-degenerate conditions. In 2003, Gao et al [13] gave a complete solution classification of the *P3P* problem using Wu's method. But the classifications are algebraic and the geometrical meaning of the algebraic conditions is not clear.

In this paper, we study one practical configuration of the *P3P* problem and gave their triangular decomposition by Wu's zero decomposition algorithm [10,11]. In practice, most manmade objects, such as building, contain symmetries. For the model shown in Figure 2, the positions of three points can be characterized using lesser parameter. This makes it possible to find the geometrical solution classifications for the *P3P* problem. Moreover, we also gave some clear geometrical conditions under which there are multi-solution for the *P3P* problem. This kind of criteria is much simpler than the algebraic one and gives some insight into the multi-solution phenomenon. On the other hand, since the manmade objects contain the configurations, this result has much practical value.

The rest of the paper is organized as follows. In Section 2, we present the zero decomposition for the equation system about the configuration. In Section 3, we present the sufficient conditions under which there are multi-solution. In Section 4, we present the conclusions.

2 The *P3P* Equation System for a Configuration

Let *P* be the center of perspective, and *A*, *B*, *C* the control points (Fig.1). Let $|PA| = X$, $|PB| = Y$, $|PC| = Z$, $\alpha = \angle BPC$, $\beta = \angle APC$, $\gamma = \angle APB$, $p = 2\cos\alpha$, $q = 2\cos\beta$, $r = 2\cos\gamma$. From triangles *PBC*, *PAC* and *PAB*, we obtain the *P3P* equation system:

$$\begin{cases} Y^2 + Z^2 - \mathrm{p}YZ - |BC|^2 = 0 \\ Z^2 + X^2 - qXZ - |AC|^2 = 0 \\ X^2 + Y^2 - \mathrm{r}XY - |AB|^2 = 0 \end{cases} \tag{1}$$

The above equation system is for variables X, Y, Z with parameters $p,\ q,\ r,\ |AB|,\ |BC|,\ |AC|$.

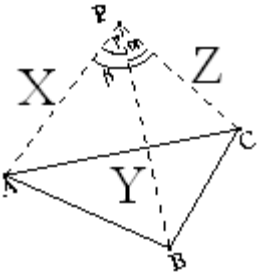


Fig. 1. The P3P problem

The solution classification for the perspective-three-point problem is the important aspect for the study of the *P3P* problem. In this section, we mainly discuss the special configuration: ABC is a right triangle that may occur quite often. In practice, there many man-made structures like building, where many right triangles exist (See Figure 2). For the configuration, we may assume that $|\mathrm{BC}|^2+|\mathrm{AC}|^2=|\mathrm{AB}|^2$.

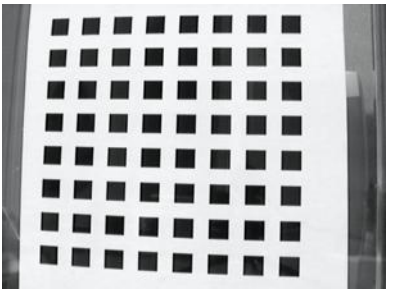


Fig. 2. An example in the practice

A set of solutions for X,Y,Z is called physical for the *P3P* problem if the following ``reality conditions" are satisfied. These conditions are assumed through out the paper.

$$\begin{cases} \mathrm{X}>0,\mathrm{Y}>0,\mathrm{Z}>0,|\mathrm{AB}|>0,|\mathrm{AC}|>0,|\mathrm{BC}|>0 \\ |\mathrm{AB}|+|AC|>|\mathrm{BC}|,|\mathrm{AC}|+|\mathrm{BC}|>|\mathrm{AB}|,|\mathrm{AB}|+|\mathrm{BC}|>|\mathrm{AC}| \\ 0<\alpha,\beta,\gamma<\pi,0<\alpha+\beta+\gamma<2\pi \\ \alpha+\beta>\gamma,\alpha+\gamma>\beta,\beta+\gamma>\alpha \\ \mathrm{I}_0=\mathrm{p}^2+\mathrm{q}^2+\mathrm{r}^2-\mathrm{pqr}-1\neq 0\ (\text{Points P, A, B, C are not co-planar [8]} \end{cases} \tag{2}$$

To simplify the equation system, set X=xZ, Y=yZ, $|AB|=Z\sqrt{v}$, $|BC|=Z\sqrt{av}$, $|AC|=Z\sqrt{bv}$ in (1). Since $Z=|PC|\neq 0$ and $|\mathrm{BC}|^2+|\mathrm{AC}|^2=|\mathrm{AB}|^2$, we obtain the following equivalent equation system:

$$\begin{cases} y^2 + 1\text{-}yp\text{-}av = 0 \\ x^2 + 1\text{-}xq\text{-}(1\text{-}a)v = 0 \\ x^2 + y^2\text{-}xyr\text{-}v = 0 \end{cases} \tag{3}$$

The equivalent correspondence is:

$$\{(x,y,v)\} \xleftarrow{|AB|=\sqrt{v}Z} \{(x,y,Z)\} \xleftarrow{X=xZ,Y=yZ} \{(X,Y,Z)\}$$

Since $|r| < 2$, we have $\mathrm{v} = \mathrm{x}^2 + \mathrm{y}^2 - \mathrm{xyr} > 0$. Thus Z can be uniquely determined by $Z = |AB| / \sqrt{v}$. Eliminating v from (1), we have

$$\begin{aligned} \mathrm{p}_1 &= (1-\mathrm{a})\mathrm{y}^2 - \mathrm{ax}^2 - \mathrm{py} + \mathrm{arxy} + 1 = 0 \\ \mathrm{p}_2 &= \mathrm{ax}^2 - (1-\mathrm{a})\mathrm{y}^2 - \mathrm{qx} + (1-\mathrm{a})\mathrm{rxy} + 1 = 0 \end{aligned} \tag{4}$$

This has the same number of physical solutions with (1). Now the *P3P* problem is reduced to finding the positive solutions of two quadratic equations (4).

3 Zero Structure for the Equation System

Wu's zero decomposition method [11, 12] is a general method to solve systems of algebraic equations. It may be used to represent the zero set of a polynomial equation system as the union of zero sets of equations in triangular form, that is, equation systems like:

$$f_1(u, x_1) = 0,\ f_2(u, x_1, x_2) = 0, \cdots f_p(u, x_1, \cdots x_p) = 0$$

Where the *u* could be considered as a set of parameters and the *x* are the variables to be determined. As shown in [11], solutions for an equation system in triangular form are well determined. For instance, the solution of an equation system in triangular form can be easily reduced to the solution of unvaried equations.

For a polynomial set *PS* and a polynomial *I*, let $\mathrm{Zero(PS)}$ be the set of solutions of the equation system $\mathrm{PS} = 0$, and $\mathrm{Zero(PS/I)} = \mathrm{Zero(PS)} - \mathrm{Zero(I)}$. Among the ``reality conditions'' listed in (2), $\mathrm{I}_0 \neq 0$ could be used to simplify the computation. Therefore, we consider $\mathrm{Zero(ES/I}_0\mathrm{)}$. Using Wu-Ritt's zero decomposition method [11, 12] to equation system (4) under the order $x \prec y$, we obtain the following characteristic set $[f_1, f_2]$:

$$\begin{cases} f_1 = -r^2ax^4 + par(r+2)x^3 + (-p^2a - qpar - 2r^2a + q^2 - qrp + r^2 - \\ q^2a)x^2 + (qp^2 + 2par - 4q + 4qa)x - p^2 + 4 - 4a = 0 \\ f_2 = (a-1)(pq-2r)^2 y - (qx-2)(-r^3ax^3 + ar^2(rq+p)x^2 + (r^3 + \\ rq^2 - 2r^3a - r^2pq - rq^2a)x + r^2p - 4rq + 4rqa + pq^2 - pq^2a = 0 \end{cases} \tag{5}$$

Since $a > 0, b = 1 - a \neq 0$, let $J = r(pq - 2r)$. We have following zero relation: $\mathrm{Zero}(4)/I_0 = \mathrm{Zero}((5)/J) \bigcup \mathrm{Zero}((4,r)/I_0) \bigcup \mathrm{Zero}((4, pq-2r)/I_0$ Consider $r = 0$ and $pq - 2r = 0$. We decompose the equation system (4) into following five disjoint components: $\mathrm{Zero}((4)/I_0) = \bigcup_{i=1}^{5} C_i$. In the above formula, $C_i = Zero(TS_i/T_i), i = 1, \cdots, 5$, where TS_i are polynomial equations in triangular form and T_i are polynomials. For corresponding *P3P* problem about the configuration, the five components are quite simple and the geometric meanings are also clear. The zero decomposition can bring some new insight into a better understanding of multiple solutions problem in the *P3P* problem, and could be used as some theoretical guide to arrange control points in real applications.

$$\begin{cases} f_1 = -r^2ax^4 + par(r+2)x^3 + (-p^2a - qpar - 2r^2a + q^2 - qrp + r^2 - \\ q^2a)x^2 + (qp^2 + 2par - 4q + 4qa)x - p^2 + 4 - 4a; \\ f_2 = (a-1)(pq-2r)^2 y - (qx-2)(-r^3ax^3 + ar^2(rq+p)x^2 + (r^3 + \\ rq^2 - 2r^3a - r^2pq - rq^2a)x + r^2p - 4rq + 4rqa + pq^2 - pq^2a \end{cases} \tag{6}$$

$$\begin{cases} q^2(a-1)y^2 + (-2rqa + 2rq)y - q^2 + 4a; \\ qx - 2 \end{cases} \tag{7}$$

$$\begin{cases} (p^2a - q^2 + q^2a)x^2 + (4q - 4qa - qp^2)x + p^2 - 4 + 4a; \\ py + qx - 2 \end{cases} \tag{8}$$

$$\begin{cases} 2x - q; \\ ry - q \end{cases} \tag{9}$$

$$\begin{cases} qx - 1; \\ py - 1 \end{cases} \tag{10}$$

The following table gives the maximal number of solutions for each component.

Table 1. The maximal number of solutions for each component

C_i	1	2	3	4	5
NO. Of solutions	4	2	2	1	1

4 Some Geometrical Conditions of Multiple Solutions for the *P3P* Problem

The multiple solutions is an important problem for the *P3P* problem. Firstly, the problem of multiple solutions is always a main concern in any real applications in the PnP problem. If the geometrical conditions of multiple solutions are available, the problem of multiple solutions can be better understood and it seems the algebraic conditions in the literature are not instructive enough. Secondly, in many real applications, if the geometrical conditions of multiple solutions are available, they can be used as theoretical guides to avoid harmful multiple solution problem by properly arranging the control points. In this section, we study two geometric configurations that occur quite often and obtain geometrically some sufficient conditions for multiple positive solutions of the corresponding *P3P* problem.

Let us assume that $a+b=1$ and $r=0$. The corresponding *P3P* equation system becomes

$$\begin{cases}(1-a)y^2-py-ax^2+1=0\\ ax^2-qx+1-(1-a)y^2=0\end{cases} \tag{11}$$

Using Wu-Ritt's method, this equation system has the following disjoint components:

$$\begin{cases}((p^2+q^2)a-q^2)x^2-(p^2+4a-4)qx+p^2+4a-4=0\\ py+qx-2=0\end{cases} \tag{12}$$

$$\begin{cases}qx-1=0\\ py-1=0\\ (p^2+q^2)a-q^2=0\end{cases} \tag{13}$$

$$\begin{cases}qx-2=0\\ q^2(a-1)y^2+4a-q^2=0\\ p=0\end{cases} \tag{14}$$

Let us assume that $a+b=1$ and $r=0$. The corresponding *P3P* equation system becomes

$$\begin{cases} x^2+py-1=0 \\ y^2+qx-1=0 \end{cases} \tag{15}$$

Using Wu-Ritt's method, this equation system has the following disjoint components:

$$\begin{cases} x^4-2x^2+p^2qx-p^2+1=0 \\ x^2+py-1=0 \end{cases} \tag{16}$$

$$\begin{cases} x-1=0 \\ y^2+q-1=0 \\ p=0 \end{cases} \tag{17}$$

From the triangle decomposition, we may have the following theorem.

Theorem 1. If $a+b=1$ and $r=0$ or $a=b=1$ and $r=0$, then the corresponding *P3P* problem has at most two physical solutions.

Proof. The Proof is clear.

Theorem 2. If $a+b=1$ and $p=r=0$ or $a=b=1$ and $p=r=0$, then the corresponding *P3P* problem have a unique physical solution.

Proof. For the geometric configuration $a+b=1, r=p=0$ and $a=b=1$, $r=p=0$, the zero component of corresponding *P3P* equation system are respectively:

$$\begin{cases} qx-2=0 \\ q^2(a-1)y^2+4a-q^2=0 \end{cases} \tag{18}$$

$$\begin{cases} x-1=0 \\ y^2+q-1=0 \end{cases} \tag{19}$$

From the practicing meaning, there is at least one solution for the *P3P* problem. So, it is clear that the corresponding *P3P* problem have unique solution from their zero decomposition. From above zero decomposition, we have the observations that the corresponding *P3P* for these geometric configurations most probably will have one solution.

5 Conclusion

In this paper, we investigate a practical configuration for the *P3P* problem and give the zero structure for the corresponding P3P equation system. The results are practical and more instructive. We also give some sufficient conditions under which there are multi-solution for the *P3P* problem. In future, the conditions for three solutions, two solutions will be investigated. In addition, the instability analysis of solutions is also an interesting topic to pursue.

References

1. Abidi, M.A., Chandra, T.: A New Efficient and Direct Solution for Pose Estimation Using Quadrangular Targets: Algorithm and Evaluation. IEEE Transaction on Pattern Analysis and Machine Intelligence 17(5), 534–538 (1995)
2. Fischler, M.A., Bolles, R.C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartomated Cartography. Communications of the ACM 24(6), 381–395 (1981)
3. Quan, L., Lan, Z.: Linear N-Point Camera Pose Determination. IEEE Transaction on Pattern Analysis and Machine Intelligence 21(8), 774–780 (1999)
4. Wolfe, W.J., Mathis, D., Weber, C., Magee, M.: The Perspective View of Three Points. IEEE Transaction on Pattern Analysis and Machine Intelligence 13(1), 66–73 (1991)
5. DeMenthon, D.F., Davis, L.S.: Exact and Approximate Solutions of the Perspective-Three-Point Problem. IEEE Transaction on Pattern Analysis and Machine Intelligence 14(11), 1100–1105 (1992)
6. DeMenthon, D.F., Davis, L.S.: Model-Based Object Pose in 25 Lines of Code. International Journal of Computer Vision 15, 123–141 (1995)
7. Su, C., Xu, Y., Li, H., Liu, S.: Necessary and Sufficient Condition of Positive Root Number of P3P Problem. Chinese Journal of Computer Sciences 21, 1084–1095 (1998) (in Chinese)
8. Su, C., Xu, Y., Li, H., Liu, S.: Application of Wu's Method in Computer Animation. In: The Fifth Int. Conf. on Computer Aided Design/Computer Graphics, vol. 1, pp. 211–215 (1997)
9. Yang Lu, A.: Simplified Algorithm for Solution Classification of the Perspective-three-point Problem. Mahematics-Mechanization Research Preprints, MMRC, CAS 17, 135–145 (1998)
10. Wentsun, W.: Basic principles of mechanical theorem-proving in elementary geometries. J. Systems Science and Mathematical Science 4(3), 207–235 (1984)
11. Wu, W.T.: Mathematics Mechanization. Science Press Beijing. English Version Kluwer Aacademic Publishers, London (2000) (in Chinese)
12. Lu, Y., Jingzhong, Z., Xiaorong, H.: Non-linear Equation System and Automated Theorem Proving, pp. 137–176. Shanghai Press of Science Technology and Education, Shanghai (1996)
13. Gao, X.S., Hou, X.R., Tang, J.L., Cheng, H.: Complete Solution Classification for the Perspective-Three-Point Problem. IEEE Transaction on Pattern Analysis and Machine Intelligence 25(8), 534–538 (2003)

Face Verification Based on AdaBoost Learning for Histogram of Gabor Phase Patterns (HGPP) Selection and Samples Synthesis with Quotient Image Method

Jianfu Chen, Xingming Zhang, and Jinsheng Li

School of Computer Science and Engineering, South China University of Technology
381#, Wushan road,Guang zhou,Guangdong, China, 510640
txjchen@gmail.com, cszxm@scut.edu.cn

Abstract. Face verification technology is widely used in public safety, e-commerce, access control, and so on. We propose a novel face verification approach, which combines a relatively new object descriptor—Histogram of Gabor Phase Patterns (HGPP), AdaBoost Algorithm selecting HGPP features and learning binary classifier, and Quotient Image method synthesizing face images under new illumination conditions. Although Gabor wavelets have been widely used in face recognition, previous studies mainly focus on the magnitude information of Gabor feature, while neglect the phase information of it. We use HGPP as an attempt to utilize the neglected Gabor phase information in face verification. Then AdaBoost algorithm trains binary classifiers, meanwhile significantly reduce the dimension of HGPP. Further, the novel strategy that synthesizes and extends training samples with Quotient Image method enhances our algorithm's robustness for illumination variation. Experiments demonstrate our novel approach is able to achieve promising face verification results under different illumination conditions.

Keywords: Face Verification, Gabor Phase, HGPP, AdaBoost, Quotient Image, Synthesize Samples, Illumination invariance.

1 Introduction

Recent years, face recognition technology (including face identification and face verification), as a biometrical technique, have received more and more concern, increasingly applied in such fields as access control, security surveillance, human-computer interaction and so forth, and bear significant theoretical meanings in pattern recognition field.

Due to its robustness to illumination variation, sensitivity to image texture, and resemblance to mammals' vision convex, Gabor features have been employed in face identification and verification and offer outstanding results, especially when addressing illumination variation problem[1][2][6][7].

However, previous studies based on Gabor wavelets mainly focus on utilizing the magnitude information of Gabor features. In fact, the phase information of Gabor features is also very discriminative, as proved by its application in iris and palmprint identification [8][9]. Recently, Zhang et al. [3] studied to utilize the phase information

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 430–437, 2008.
© Springer-Verlag Berlin Heidelberg 2008

of Gabor features in face identification, proposing Histogram of Gabor Phase Patterns (HGPP) and achieved striking result in face identification.

We combine HGPP features proposed by Zhang et al. [3] and AdaBoost learning algorithm to study face verification and obtain promising results in our experiments, as an attempt to utilize the phase information of Gabor features in face verification fields. In [3], HGPP is applied in face identification—find several most similar persons for a probe image from a database, and the dimension of HGPP is quite high. We introduce HGPP into face verification--that is--verifying whether a probe image belongs to the claimed registered person. AdaBoost is a simple yet effective stage-wise binary classifier learning algorithm, performing feature selection and nonlinear learning at the same time [4], and has been successfully applied in face detection[10] and other fields[4][5]. In our algorithm, AdaBoost algorithm trains a binary classifier for face verification, meanwhile selects and thus reduces the dimension of HGPP features.

In addition, in order to further enhance the robustness to illumination variation and help to solve small-sample-size problem, we proposed a novel approach to synthesize image samples under new illumination conditions for registered images. Quotient Image method, proposed by Shashua et al.[11], can generate an illumination invariant signature image (named Quotient Image) for a given image, and the obtained Quotient Image could be used to render images under new illumination conditions for a given image. We use this method to synthesize and extend face samples.

Therefore, the face verification approach proposed by us has the following features: (1) utilize HGPP that is newly introduced in face identification [3] to study face verification, as an attempt to apply the previously neglected phase information of Gabor features in face verification. (2) apply AdaBoost algorithm to train binary classifiers, and significantly reduce the dimension of HGPP at the same time. (3) The novel approach that synthesize and extend face samples based on Quotient Image method strengthen our algorithm's robustness to illumination variation, and is conducive to solve small-sample-size problem which usually exists in practice.

The rest of this paper is organized as follows: section 2 sketches out the overall flow of our proposed algorithm; section 3 introduces HGPP feature; section 4 details the approach of training binary classifier based on AdaBoost algorithm; section 5 describes the novel Quotient Image based approach to synthesize face samples; section 6 and 7 is experiments and discussions, and conclusion respectively.

2 Overall Flow of Algorithm

The overall flow of proposed algorithm, which uses AdaBoost algorithm to select HGPP features and learn binary classifiers, and synthesize, extend face samples based on Quotient Image method, is shown in Fig.1. The training flow is indicated by real line arrows, while the testing/verification flow by dash line arrows. We train a binary classifier for each registered person for face verification. The details of the algorithm will be illustrated in section 3 to 5.

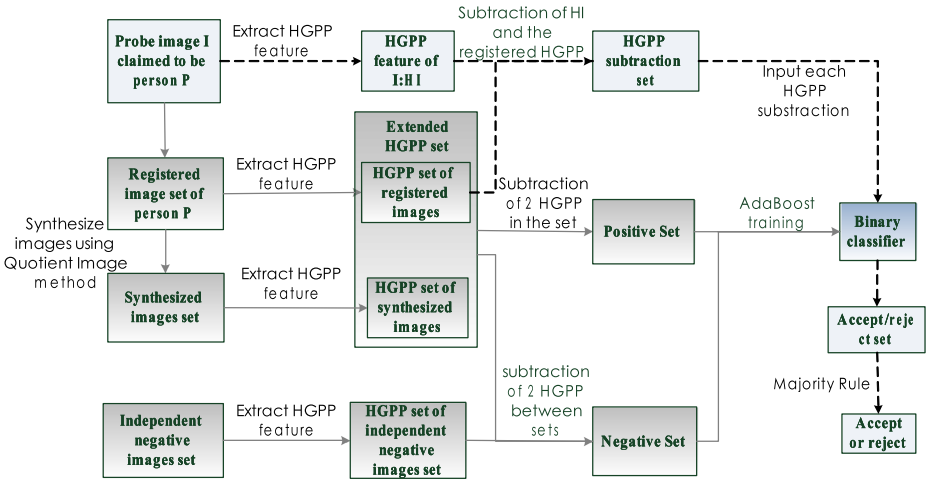


Fig. 1. Overall Flow of Proposed Face Verification Algorithm

3 HGPP: Histogram of Gabor Phase Patterns

Gabor wavelets is actually the result of Gaussian function horizontal movement in frequency domain, and have been employed in face recognition and offer outstanding results [1][2][6][7]. We use the same method and parameters as [2] to extract Gabor features.

However, former studies on face recognition have mainly explored the magnitude information of Gabor feature while neglected its phase information which is proved to be also quite discriminative. Zhang et al. [3] proposed Histogram of Gabor Phase Patterns(HGPP), which encode the phase information of Gabor feature, into face identification and achieve striking discrimination ability and illumination invariance. We introduce HGPP into face verification here.

The final HGPP feature of an image i is a series of histograms:

$$HGPP_i = [h_{i,1}, h_{i,2}, ..h_{i,Q}] \tag{1}$$

where $h_{i,j}$ is a histogram, j=1, … Q, and Q is the number of histograms.

4 Using AdaBoost Algorithm to Learn Binary Classifiers Based on HGPP

AdaBoost is a simple yet effective stage-wise binary classifier learning algorithm, which constructs strong classifier from a series of easily learnable weak classifiers whose correct classification rate are needed to be just a little greater than 50% [4]. It is applicable to the binary classification problem of face verification, and can achieve the effect of selecting and thus reducing the dimension of HGPP features at the same time. The basic flow of AdaBoost algorithm is as follows [4].

In our algorithm, each training sample for AdaBoost learning is generated by subtraction of 2 HGPP features of 2 face images. If the 2 face images which generate the sample belong to the same person, the generated sample is positive sample; and vice versa. The subtraction of 2 HGPP features is defined as bellow:

The subtraction of 2 HGPP features: $HGPP_i = [h_{i,1}, h_{i,2}, \ldots, h_{i,Q}]$ and $HGPP_j = [h_{j,1}, h_{j,2}, \ldots, h_{j,Q}]$ *is* a vector formed by the Chi distance between each pair of corresponding histograms in the 2 series of histograms, that is:

$$HGPP_i - HGPP_j = [CHI(h_{i,1}, h_{j,1}) \;\; CHI(h_{i,2}, h_{j,2}) \;\; \ldots \;\; CHI(h_{i,Q}, h_{j,Q})] \tag{2}$$

where CHI is a method of calculating dissimilarity between 2 histograms, the Chi distance between 2 histograms is defined as [5]:

$$CHI(S, M) = \sum_{i=1}^{n} \frac{(S_i - M_i)^2}{S_i + M_i} \tag{3}$$

where n is the bin number of histogram.

In our experiments, we set AdaBoost training iteration times: T = 200. Each iteration train threshold functions as weak classifiers on each dimension of the subtraction of HGPP features. Hence, we need at most 200 (the dimensions upon which the weak classifiers are selected out in different iterations, may be duplicate) histograms for face verification, while the typical original HGPP feature is 5760 32-scale histograms. Therefore, we implement considerable dimension reduction of HGPP feature at the same time.

The procedure of applying the binary classifier trained by AdaBoost algorithm to face verification is as the flow indicated by the dash line arrows in fig.1.: for a given probe face image I, and the claimed person identity P, extract HGPP feature of image I: HI, subtract HI from each HGPP feature in the HGPP feature set of person P's registered images and input the difference to binary classifier **D_p, and gain a judgment true/accept or false/reject.** We do above judgments R times (R is size of **HGPP set of registered images**). If more than half of the judgment results is true/accept, we accept the person corresponding to I is person P; and vice versa.

5 Synthesizing Face Samples under New Illumination Conditions Based on Quotient Image Method

5.1 Quotient Image Method

Shashua et al. [11] propose quotient image method based on Lambertian Model. The quotient image Q_y of face y against face a is defined by the ratio of the albedo (surface texture) of face y to the albedo (surface texture) of face a. So the quotient image Q_y depends only on the relative surface texture information and thus is independent of illumination.

And we can use quotient image to generate new face images under different illumination conditions. In practice, quotient image method use a bootstrap set of 3N (N is the number of the persons) pictures taken from three fixed (linearly independent)

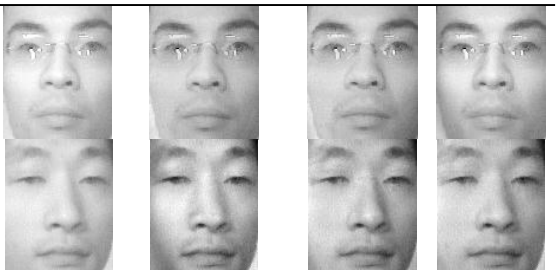
light sources s1, s2, s3 (the light sources are not known) to generate quotient image of face image y against the bootstrap.

5.2 Synthesizing Training Image Samples

The motivation of synthesizing training image samples comes from the results of our comparative experiments (see section 6). From the experiment results, we assume that the requirement of subtraction between HGPP features to generate training samples is likely to cause our algorithm to tend to be affected by illumination variation, in that the over fitting phenomena may exists in the AdaBoost training procedure when the training samples are from relatively single type of illumination conditions.

To address the above problem, we utilize Quotient Image method to render new images under random illumination conditions for the images in the registered images set $F = [f_1, f_2, ... f_P]$ of each person. The rendered images join and extend the registered images set F, thus we gain a larger image sample set to generate training samples for AdaBoost learning (See fig.1. and table.1.). And experiments demonstrate that this approach improve our algorithm's robustness to illumination variation.

Table 1. examples of using Quotient Image method to synthesize images under new illumination conditions

Original Images		Original Images	
			

6 Experiments and Discussions

We conduct the experiments on the "UCity" face database collected by our laboratory, which includes face images of 44 persons under 3 different illumination conditions (denoted as illumination A, B, and C). 40 images with pose variation are collected for each person under every illumination condition. All the face images used in experiments was cropped to 92 pixels by 112 pixels. The samples of the database are shown in table 2.

We first generate independent HGPP set $Neg = [N_1, N_2, ..., N_N]$ which will be used to generate negative samples for AdaBoost training for each registered person.Choose 20 persons with 4 images each in UCity database to extract HGPP features as Neg, thus N is 20*4=80. And select other images of the above 20 persons to constitute the Bootstrap Set used in Quotient Image method [11]: select one image of each person

Table 2. The samples of Ucity Face Database

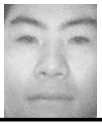
Illumination A						
Illumination B						
Illumination C						

Table 3. Experiments Results

Methods	Illumination Condition of registered Pictures.	Average FRR In Set A (%)	Average FRR In Set B (%)	Average FRR In Set C (%)	Average FRR In All Sets (%)
HGPP+AdaBoost+NoSyn*	A	5.5	21.45	25.00	17.32
HGPP+AdaBoost+NoSyn*	ABC	3.47	7.21	8.72	6.47
HGPP+AdaBoost+Syn*	A	2.40	11.60	12.00	8.67
FLD+SVM	A	4.75	22.21	37.52	21.49

*NoSyn denotes the strategy of synthesizing image samples based on Quotient Image method is not used, Syn denotes the synthesizing strategy is used.

under each illumination condition. The remaining 24 persons in the database are as registered persons and play as impostors each other in testing.

The experiments can be divided into 2 groups by whether they use the strategy of synthesizing image samples based on Quotient Image method.

1) When the strategy of synthesizing image samples is not used, 15 images are registered for each person P. In order to evaluate the algorithm's robustness to illumination variation, we compare the experiment results when the 15 registered images of each person P are from a single illumination and from three illumination conditions with 5 images under each illumination condition.
2) When the strategy of synthesizing image samples is used, 9 images are registered under illumination A for each person, from which 5 images are selected to be rendered by Quotient Image method under 3 random illumination conditions. Hence, 15 synthesized images are gained; they and the 9 registered images (together 24 images) make up the extended registered images set upon which the binary classifier training are conducted. (see fig.1.)

For each of the two strategies, we train a binary classifier for each person P. All the face images of P except the training images are used to test FRR, and the face images of all the other 23 persons are used to test FAR. For comparison with traditional algorithm, we at the same time test FLD+SVM method [12]. Adjust parameters and make FAR is 0.1%, calculate the average FRR of all persons' binary classifiers, the result of experiments are as Table 3.

Comparing the first and second result row of Table 3, when the strategy of synthesizing images samples is not used, the error rate is rather lower when the registered images for each person are from three type of illumination than from only illumination A. We analyze the reason for this distinction: from experiment data, we discover that when the training images are only from one type of illumination condition, due to the training samples for AdaBoost learning are generated by the subtraction of HGPP features, sometimes our AdaBoost learning only choose the weak classifiers which are on the same specific dimension of training samples to construct strong classifiers, for the error rates of these weak classifiers which are on the same specific dimension are 0. But our experiments show that these sort of final binary classifiers have poor generalization ability under other illumination conditions than the one training images are from.

Although collecting training images from multiple illumination conditions gives quite good verification rate, in practical applications collecting just a small number of training samples under single type of illumination would be favorable, because usually it is hard to get enough training samples. So we apply the strategy of synthesizing training samples under new illumination conditions. And experiments (see Table 3) indicate that the error rate of synthesis strategy is a little higher than that of the one when original training samples are from three illumination conditions, but is significantly lower in comparison to that of the strategy when training samples are from only one type of illumination and no images are synthesized. This fact demonstrate that our training images samples synthesis strategy effectively circumvent the overfitting problem of AdaBoost training when training samples are from a single type of illumination, and at the same time we use less training samples. Further, comparing to FLD+SVM method, our algorithm is no doubt superior.

7 Conclusion

We propose a novel face verification approach, which combines a relatively new object descriptor—Histogram of Gabor Phase Patterns (HGPP), AdaBoost Algorithm selecting HGPP features and learning binary classifier, and Quotient Image method synthesizing face images under new illumination conditions.

Our approach utilizes HGPP that is newly introduced in face identification [3] to study face verification, as an attempt to apply the previously neglected phase information of Gabor features in face verification.

We apply AdaBoost algorithm not only to train binary classifiers, but also significantly reduce the dimension of HGPP at the same time. However, experiments show that due to the requirement of subtraction of HGPP features of training image samples to generate its training samples, AdaBoost learning may over-fit on the training

images samples when they are from a single type of illumination, and thus lack of generalization ability under other illumination conditions.

To solve the above problem, the novel approach that synthesize and extend face samples under different illumination conditions based on Quotient Image method strengthen our algorithm's robustness to illumination variation, and is conducive to solve small-sample-size problem which usually exists in practice. Experiments demonstrate that even if the training image samples are from a single type of illumination, our algorithm is able to achieve promising results under different testing illumination conditions.

References

1. Shen, L., Bai, L.: A review on Gabor wavelets for face recognition. Pattern Anal. Applic. 9, 273–292 (2006)
2. Liu, D., Lam, K., Shen, L.: Optimal sampling of Gabor features for face recognition. Pattern Recognition Letters 25, 267–276 (2004)
3. Zhang, B., Shan, S., Chen, X., Gao, W.: Histogram of Gabor Phase atterns (HGPP): A Novel Object Representation Approach for Face Recognition. IEEE Trans. Imag. Proc. 16(1) (2007)
4. Zhang, L., Li, S., Qu, Z., Huang, X.: Boosting Local Feature Based Classifiers for Face Recognition. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2004) (2004)
5. Zhang, G., Huang, X., Li, S., Wang, Y., Wu, X.: Boosting Local Binary Pattern(LBP)-Based Face Recognition. In: Li, S.Z., Lai, J.-H., Tan, T., Feng, G.-C., Wang, Y. (eds.) SINOBIOMETRICS 2004. LNCS, vol. 3338, pp. 179–186. Springer, Heidelberg (2004)
6. Wiskott, L., et al.: Face recognition by elastic bunch graph matching. IEEE Trans. PAMI 19(7), 775–779 (1997)
7. Shen, L., Bai, L., Fairhurst, F.: Gabor wavelets and General Discriminant Analysis for face identification and verification. Image and Vision Computing (2006)
8. Daugman, J.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. PAMI 15(11), 1148–1161 (1993)
9. Zhang, D., Kong, W., You, J., Wong, M.: Online palmprint identification. IEEE Trans. PAMI 25(9), 1041–1050 (2003)
10. Viola, P., Jones, M.: Robust real time object detection. In: IEEE ICCV Workshop on Statistical and Computational Theories of Vision, Vancouver, Canada, July 13 (2001)
11. Shashua, A., Riklin-Raviv, T.: The Quotient Image: Class-Based Re-Rendering and Recognition with Varying Illuminations. IEEE Trans PAMI 23(2) (February 2001)
12. Zhang, X., Li, H.: A Face Verification Algorithm Based on Negative Independent Sample Set and SVM. Journal of Computer Research and Development, 2138–2143 (2006) (in Chinese)

Predicting Epileptic Seizure by Recurrence Quantification Analysis of Single-Channel EEG

Tianqiao Zhu[1], Liyu Huang[1], Shouhong Zhang[1], and Yuangui Huang[2]

[1] School of Electronic Engineering, Xidian University, Xi'an, 710071, China
[2] Xijing Hospital, The Fourth Military Medical University, Xi'an, 710032, China
huangly@mail.xidian.edu.cn,
tqzhu@mail.xidian.edu.cn

Abstract. This paper presents a promising approach, based on the combination of two different approaches, including recurrence quantification analysis(RQA) of single-channel EEG and artificial neural network (ANN), which predicts seizure onsets in 17 consenting patients with generalized epilepsy. Eight channels of EEG were collected in each patient in Epilepsy Center of Xijing Hospital. The seven measures were extracted from the raw EEGs by RQA, and then a four layer(7-6-2-1) ANN was employed for prediction, using the measures as the inputs. The performance obtained for the proposed scheme in predicting succedent seizures is: sensitivity 50~72.2%, specificity 61.9~73.8% and accuracy 58.3~70%, depending on the different EEG leads.

1 Introduction

The incidence of epilepsy is estimated to be about 0.6 ~ 1% in the general population. The main hallmark of epilepsy is recurrent seizures[1]. Moreover, the disease can appear at any age[2, 3]. A system able to herald seizures far enough in advance to allow preventive action would reduce morbidity and mortality, and greatly improve the quality of life for epilepsy patients.

In fact, numerous efforts have been made to develop processed EEG derivatives for seizure predication[2-5]. Nevertheless, none of these technologies has been showed to be sufficiently reliable for routine use for predicting seizures up to now.

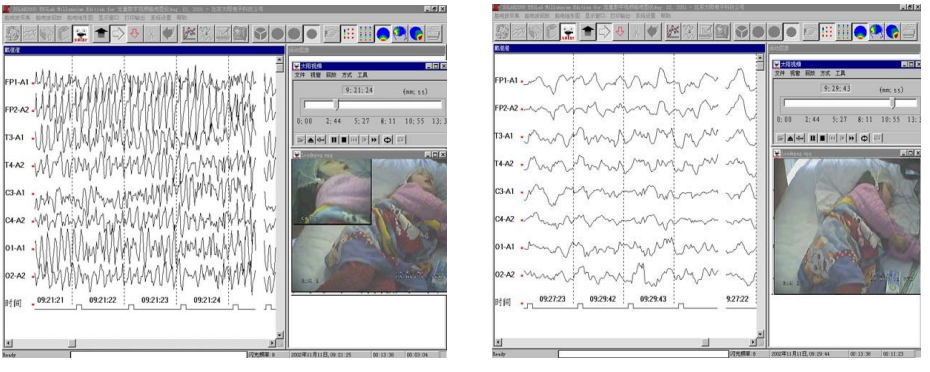
In this paper, we report a new approach for seizure prediction, which is based on the combination of recurrence quantification analysis of single-channel scalp EEG and ANN.

2 Materials and Data Acquisition

Our EEG data were chosen from routinely acquired EEG data from patients with generalized epilepsy in Epilepsy Center of Xijing Hospital. The EEG was monitored using a QDBS1018V monitor (Beijing solar tech Ltd., China), which stores continuous EEG data on compact disc along with video images. Synchronization of video

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 438–445, 2008.
© Springer-Verlag Berlin Heidelberg 2008

and EEG was achieved and stored for offline analysis of clinical onsets and overall patient behavior during the hospital stay. Both viewing the video and looking at the EEG identified the patient's state of consciousness(see figure 1). Eight channels of EEG were recorded using Ag/AgCl electrodes (Fp1, Fp2, O1, O2, T3, T4, C3 and C4, they were placed in accordance with the international standard 10~20 system, with the reference points at left earlobe, A1 and right earlobe, A2), including Fp1-A1, Fp2-A2, O1-A1, O2-A2, T3-A1, T4-A2, C3-A1, C4-A2. The impedance on each of the EEG channel was required to be less than $2K\Omega$ at the start of each monitoring. The EEG data were sampled at 200*Hz*, digitized to 12bits and filtered with a high-pass filter at 0.4*Hz* and a low-pass filter at 70*Hz*. A 50*Hz* notch filter was also employed. Figure 1 is an example of one patient being monitored and the EEG and video being recorded. Figure 2 shows the EEGs of a patient before, during and after a seizure.



(a) (b)

Fig. 1. Examples of one patient being monitored during(a) and after(b) a seizure

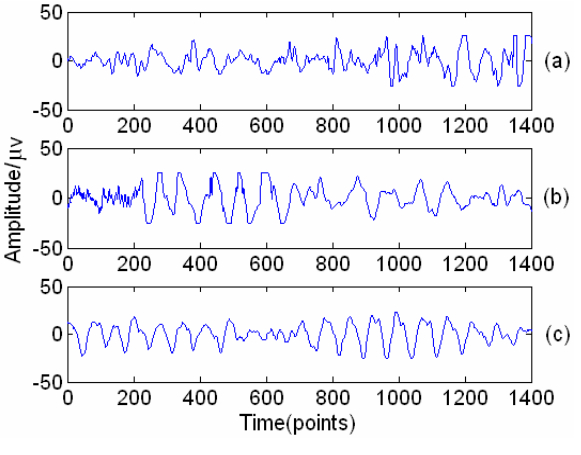


Fig. 2. EEGs of a patient before(a), during(b) and after(c) a seizure

From 17 long (more than 12 hours) clinical monitoring records of the patients, there are 60 distinct EEG epochs to be chosen to study, including 18 seizure epochs, which are selected from recordings of 60s prior to seizure onset, 42 non-seizure epochs coming from epileptic patients without seizure. As post-epileptic activity (after the onset of the seizure) can last up to an hour or so, the time from the closest a seizure onset of the non-seizure epochs must be more than one hours. Each epoch is 7s long (1,400 points).

3 Method

Figure 3 is the flow diagram of proposed system. The input to the ANN will be the 7 measure values, the output will be result of the prediction. We describe these procedures briefly as follows.

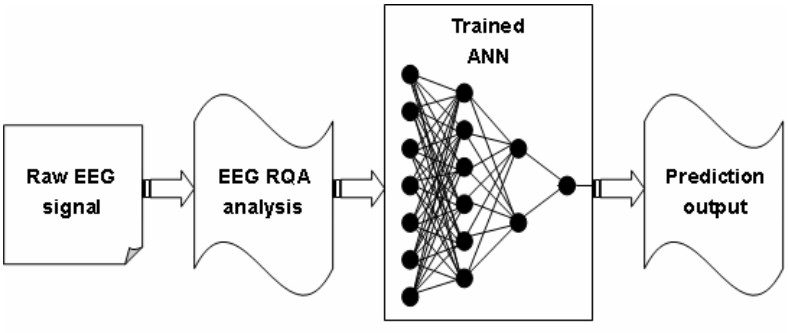


Fig. 3. System for predicting seizure using RQA of single channel EEG and ANN

3.1 Recurrence Plot

Suppose that a dynamic system can be relaxed to an attractor which has an invariant distribution in a phase space. The phase space can be constructed by expanding a one-dimensional time series u_i (e.g. EEG, from observations) into a vector set using the Taken's time delay theorem[6],

$$\vec{x}_i = (u_i, u_{i+\tau}, \cdots, u_{i+(m-1)\tau}) \tag{1}$$

the dimension *m* can be estimated with the method of false nearest neighbors [7, 8], the time delay τ can be estimated with the method of mutual information[9].

Based on the reconstruction of phase space mentioned above, Eckmann *et al.* first proposed a new method, recurrence plot (RP), to visualize the time dependent behavior of the dynamics of systems, which can be pictured as a trajectory $\vec{x}_i \in \Re^n (i = 1, \cdots, N)$ in the *n*-dimensional phase space[10]. Mathematically speaking, RP can be obtained by the calculation of the $N \times N$ matrix,

$$R_{i,j} := \Theta(\varepsilon_i - \left\| \vec{x}_i - \vec{x}_j \right\|), \quad i, j = 1, \cdots, N \tag{2}$$

where $\left\| \cdot \right\|$ is a norm (e.g., the Euclidean norm), ε_i is a cutoff distance, and $\Theta(x)$ is the Heaviside function($\Theta(x) = 1$, if $x > 0$; $\Theta(x) = 0$, if $x \le 0$).

The cutoff distance ε_i defines a sphere centered at $\vec{x}_i$. If $\vec{x}_j$ falls within this sphere, the state will be close to $\vec{x}_i$ and thus $R_{i,j} = 1$. ε_i can be either constant for all $\vec{x}_i$ or they can vary in such a way that the sphere contains a predefined number of close states. In this paper the Euclidean norm is used and ε_i is determined by 5% of the maximum value of all computed norm values. The binary values in $R_{i,j}$ can be simply visualized by a matrix plot with the colors black (1) and white(0).

The recurrence plot exhibits characteristic large-scale and small-scale patterns that are caused by typical dynamical behavior [10,11], e.g., the diagonals mean similar local evolution of different parts of the trajectory; the horizontal or vertical black lines mean that state does not change for some times.

3.2 Recurrence Quantification Analysis

Zbilut and Webber have presented the recurrence quantification analysis (RQA) to quantify an RP [11], and Marwan et al. extend the definition on the vertical structures[12]. Seven RQA variables are usually examined: (1) *Recurrence rate(RR)*, quantifies a percentage of the plot occupied by recurrent points; (2) *Determinism (DET)*, quantified a percentage between the recurrent points that form upward diagonal line segments and the entire set of recurrence points. The diagonal line consists of two or more points that are diagonally adjacent with no intervening white space; (3)*Laminarity(LAM)*, quantified a percentage between the recurrence points forming the vertical line structures and the entire set of recurrence points; (4)*L-Mean*: quantified an average of the lengthes of upward diagonal lines, this measure contains information about the amount and the length of the diagonal structures in the RP; (5)*Trapping time(TT)*: quantified an average of the lengthes of vertical line structures; (6)*L-Entropy(L-ENTR)*, is from Shannon's information theory, quantified a Shannon entropy of diagonal line segment distributions; (7)*V-Entropy (V-ENTR)*, quantified a Shannon entropy of vertical line segment distributions.

3.3 Artificial Neural Networks(ANN)

A four-layer ANN（7-6-2-1） was used in our system. We build up the network in such a way that each layer is fully connected to the next layer.

Training of the ANN, essentially an adjustment of the weights, was carried out on the training set, using the back-propagation algorithm. This iterative gradient algorithm is designed to minimize the root mean squared error between the actual output and the desired output.

The ANN was trained using 'leave one out' method, because of our small number of sample recording [13].

4 Results

The RP of the EEG recorded was computed and mapped for each patient or subject with a time series of 1400 points. Fig 4 shows the RPs of two EEG channels of a patient in different seizure stages, the plots are from Fp1-A1 EEG, 60s before seizure onset(a), C3-A1 EEG, also 60s before seizure onset(b), Fp1-A1 EEG without seizure(c) and C3-A1 EEG without seizure(d), respectively. The seven measures are extracted from these plots and used as input to the ANN for training and test.

The three measures, including accuracy, sensitivity and specificity, were used as evaluation criteria to assess the performance of our proposed scheme.

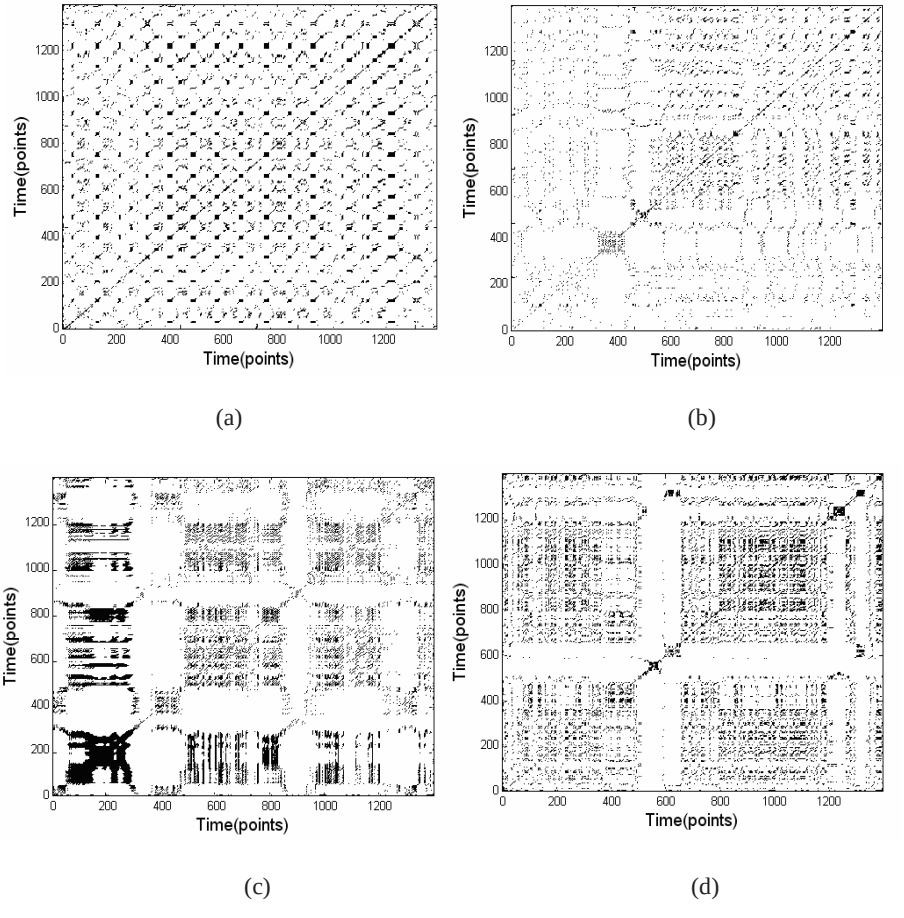


Fig. 4. RP examples of a patient in different seizure stages and EEG channels, (a) Fp1-A1 EEG, 60s before seizure onset; (b) C3-A1 EEG, 60s before seizure onset; (c) Fp1-A1 EEG without seizure; (d) C3-A1 EEG without seizure

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{3}$$

$$sensitivity = \frac{TP}{TP + FN} \times 100\% \tag{4}$$

$$specificity = \frac{TN}{TN + FP} \times 100\% \tag{5}$$

where TP is the number of true positive, TN is the number of true negative, FP is the number of false positives, FN is the number of false negatives.

The results of testing our proposed system are shown in Table 1. We illustrate the results by Fp1-A1 EEG channel in the table. In total, there are 18 seizure recordings and 42 non-seizure recordings for training and test. Nine seizure states are misclassified as non-seizure, and thirteen non-seizure states are misclassified as seizure. For seizure prediction, the average accuracy is 63.3% by using one channel EEG, Fp1-A1.

5 Discussion

In humans, epilepsy is the second most common neurological disorder, next to stroke. Up to now there is only limited knowledge about seizure-generating mechanisms. One of the most disabling features of epilepsy is the seemingly unpredictable nature of seizures. Forewarning would allow the patient to interrupt hazardous activity, lie down in a quiet place, undergo the seizure, and then return to normal activity. On the other hand, about two-thirds of affected individuals have epileptic seizures that are controlled by antiepileptic drugs or surgery, the remaining subgroup(about one-fourths of the total epilepsy patients) cannot be properly controlled by any available therapy. At present, prediction of the occurrence of the seizures from the scalp EEG, with the combination of electrical stimulation or biofeedback, is one of the most commonly researched method to reduce the frequency of epileptic seizures. No doubt the techniques capable of reliably predicting epileptic seizures would be of great value. However, if using multi-channel EEGs to achieve this aim, it is inconvenient or impractical for many application circumstances, such as implantable device of electrical stimulation or drug release system. Thus, there is currently growing interest in realizing seizure prediction from one or two-channel scalp EEG. In this paper, the method proposed is based on single channel EEG analysis.

Epileptic patients always exhibit abnormal, synchronized discharges in their brains. Although abnormal discharges are always present during seizures, the presence of the discharge patterns does not indicate the occurrence of a seizure. This makes seizure prediction so difficult.

Table 1. Test results of the proposed scheme, which only uses one-channel EEG

Analyzed lead	Epochs of different EEG states	Prediction error	Sensitivity, %	Specificity, %	Accuracy, %
Fp1-A1	60s before onset (18) without seizure (42)	*FP* = 13 *FN* = 9	50.0	69.0	63.3
Fp2-A2	60s before onset (18) without seizure (42)	*FP* = 14 *FN* = 7	61.1	66.7	65.0
T3-A1	60s before onset (18) without seizure (42)	*FP* = 16 *FN* = 9	50.0	61.9	58.3
T4-A1	60s before onset (18) without seizure (42)	*FP* = 15 *FN* = 8	55.6	64.3	61.7
C3-A1	60s before onset (18) without seizure (42)	*FP* = 13 *FN* = 7	61.1	69.0	66.7
C4-A2	60s before onset (18) without seizure (42)	*FP* = 12 *FN* = 6	66.7	71.4	70.0
O1-A1	60s before onset (18) without seizure (42)	*FP* = 15 *FN* = 5	72.2	64.3	66.7
O2-A2	60s before onset (18) without seizure (42)	*FP* = 11 *FN* = 8	55.6	73.8	68.3

Over the past twenty years, researchers have devoted their efforts to identifying algorithms that are able to predict or rapidly detect the onset of seizure activity, but none of these methods has sufficient sensitivity and specificity. The previous works indicate that both linear and nonlinear EEG analysis techniques can predict seizure onset at certain accuracy, further improvement could probably be achieved by a combined use of these techniques. In this paper, we report our works on seizure prediction, which is based on the combination of recurrence quantification analysis of single-channel scalp EEG and ANN. The results show the combination of these different methods has a certain advantage.

The EEG, quite likely governed by nonlinear dynamics like many other biological phenomena [14] and epileptic EEG possesses chaotic and fractal natures at least to some extent[15]. Nonlinear approaches to the analysis of the brain can generate new clinical measures, such as measures of RQA, as well as new ways of interpreting brain electrical function, particularly with regard to epileptic brain states. The RP analysis exhibits characteristic large-scale and small-scale patterns that are caused by typical dynamical behavior, e.g., diagonals (similar local evolution of different parts of the trajectory) or horizontal and vertical black lines (state does not change for some time)[10]. RQA is able to characterize complex biological processes, easily detecting and quantifying, too, intrinsic state changes, so overcoming the need for series stationarity and providing quantitative descriptors of signal complexity [11,12].

Nevertheless, the work reported here is preliminary. The next step in this research is to perform this analysis over a large number of patients with various epilepsies,

including partial and generalized, to test the effectiveness of the method with sufficient training and test data. The strategy in how to choose the optimal EEG channel will be another important research direction.

Acknowledgment. This work was supported by National Natural Science Foundation of China under grant NO. 60371023.

References

1. Engel Jr., J., Pedley, T.A.: Epilepsy: A Comprehensive Textbook. Lippincott Raven, Philadelphia (1997)
2. Litt, B., Echauz, J.: Prediction of Epileptic Seizures. The Lancet Neurology 1, 22–31 (2002)
3. Iasemidis, L.D.: Epileptic Seizure Prediction and Control. IEEE Trans. on Biomed.l Eng. 50(5), 549–556 (2003)
4. Lehnertz, K., Mormann, F., Kreuz, T.: Seizure Prediction by Nonlinear EEG Analysis. IEEE Engineering in Medicine and Biology Magazine, 57–64 (January/February 2003)
5. Babloyantz, A., Destexhe, A.: Low Dimensional Chaos in an Instance of Epilepsy. Proc. Nat. Acad. Sci. USA 83, 3513–3517 (1986)
6. Takens, F.: Determing Strang Attractors in Turbulence. Lecture notes in math. 898, 361–381 (1981)
7. Cao, L.: Practical Method for Determining the Minimum Embedding Dimension of a Scalar Time Series. Physica D 110, 43–50 (1997)
8. Kennel, M.B.: Determining Embedding Dimension for Phase-space Reconstruction Using a Geometrical Construction. Physical Review A 65, 3403–3411 (1992)
9. Fraser, A.M., Swinney, H.L.: Independent Coordinates for Strange Attractors from Mutual Information. Phys. Rev. A 33, 1134–1140 (1986)
10. Eckmann, J.P., Kamphorst, S.O., Ruelle, D.: Recurrence Plots of Dynamical Systems. Europhys. Lett. 4, 973–977 (1987)
11. Trulla, L.L., Giuliani, A., Zbilut, J.P., Webber Jr., C.L.: Recurrence Quantification Analysis of the Logistic Equation with Transients. Phys. Lett. A 223, 255–260 (1996)
12. Marwan, N., Wessel, N., Meyerfeldt, U.: Recurrence-plot-based Measures of Complexity and their Application to Heart-rate-variability Data. Phys. Reviw. E 66, 026702 (2002)
13. Fukunaga, K.: Introduction to Statistical Pattern Recognition, 2nd edn. Academic, San Diego (1990)
14. Lehnertz, K., Elger, C.E.: Can Epileptic Seizures Be Predicted? Evidence from Nonlinear Time Series Analysis of Brain Electrical Activity. Physical Review Letters 80, 5019–5022 (1998)
15. Yaylali, I., Kocak, H., Jayakar, P.: Detection of Seizures from Small Samples Using Nonlinear Dynamic System Theory. IEEE trans. on Biomed. Eng. 43, 743–751 (1996)

Double Sides 2DPCA for Face Recognition

Chong Lu[1,2,3], Wanquan Liu[2], Xiaodong Liu[1] , and Senjian An[2]

[1] School of Electronic and Information Engineering Dalian University of technology, China 116024
[2] Curtin University of Technology, Australia,WA 6102
[3] Dept. of Computer, YiLi Normal College, China 835000
cluxjyn@gmail.com, W.Liu@curtin.edu.au

Abstract. Recently, many approaches of face recognition have been proposed due to its broad applications. The generalized low rank approximation of matrices(GLRAM) was proposed in [1], and a necessary condition for the solution of GLRAM was presented in [2]. In addition to all these developments, the Two-Dimensional Principal Component Analysis (2DPCA) model is proposed and proved to be an efficient approach for face recognition [5]. In this paper, we proposed Double Sides 2DPCA algorithm via investigating the 2DPCA algorithm and GLRAM algorithm, experiments showed that the Double Sides 2DPCA's performance is as good as 2DPCA's and GLRAM's. Furthermore, the computation cost of recognition is less than 2DPCA and the computation speed is faster than that for GLRAM. Further, we present a new constructive method for incrementally adding observation to the existing eigen-space model for Double Sides 2DPCA, called incremental doubleside 2DPCA. An explicit formula for such incremental learning is derived. In order to illustrate the effectiveness of the proposed approach, we performed some typical experiments.

1 Introduction

Research into face recognition has flourished in recent years due to the increasing need for robust surveillance and has attracted a multidisciplinary research effort, in particular, for techniques based on PCA [6], [7], [8], [9], [10]. Recently, Ye [1] proposed the generalized low rank approximation of matrices(GLRAM) , and a necessary condition of the GLRAM was given in [2]. Jian Yang et al. [5] presented a new approach, called 2DPCA and have applied it to appearance-based face representation and recognition. With this approach, an image covariance matrix can be constructed directly using the original image matrices and this overcomes the weaknesses of PCA in which 2D face image matrices must be previously transformed into 1D image vectors. As a result, 2DPCA has two important advantages over PCA. First, it is easier to evaluate the covariance matrix accurately. Second, less time is required to compute the corresponding eigenvectors. Further, the performance of 2DPCA is usually better than PCA as demonstrated in Yang [5]. Peter M. Hall et al. Hall [3] proposed "incremental eigen analysis" and pointed out that incremental methods are important to the computer vision

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 446–459, 2008.
© Springer-Verlag Berlin Heidelberg 2008

community because of the opportunities they offer. To overcome recomputing the whole decomposition from scratch, several methods have been proposed that allow for an incremental computation of eigen images [11],[12]. Matej Artač et al. [4] used that incremental PCA for on-line visual learning and recognition. In this paper, we proposed a doubleside 2DPCA algorithm and intend to develop an eigen space update algorithm based on the doubleside 2DPCA here and use it for face recognition. We also compare the results with GLRAM[1].

An incremental method computes an eigen space model by successively updating an earlier model as new observations become available. The advantage of the incremental 2DPCA method is that it allows the construction of eigen models via a procedure that uses less storage and could be used as part of a learning system in which observations are added to the existing eigen model iteratively, the doubleside incremental 2DPCA is using incremental 2DPCA on double sides.

In our experiments we are using the AMP face image database, ORL face image database, Yale face image and Yale B face image databases [14] and the experiments showed that the doubleside 2DPCA's performance is comparable with 2DPCA's and GLRAM's. Moreover, the computation cost of recognition is less than 2DPCA and GLRAM. We also show that the resulting model is comparable in performance to the model computed with the batch method. However, the incremental model can easily be modified with incoming data and save much time in recognizing process.

This paper is organized as follows. In section 2, we introduce GLRAM and a necessary condition for optimality of GLRAM. We introduce the 2DPCA in section 3. In section 4, we propose a doubleside 2DPCA. We introduce incremental 2DPCA methods and then develop our novel approach in section 6. We present the results of the experiments which show the advantage of the proposed approach in section 7 and conclude the paper in section 8.

2 Generalized Low Rank Approximation of Matrices

The traditional low rank approximation of matrices was formulated as follows. Given $A \in \mathbb{R}^{n\times m}$, find a matrix $B \in \mathbb{R}^{n\times m}$ with rank$(B)= k$, such that

$$B = arg \min_{\mathrm{r}ank(B)=k} ||A - B||_F \tag{1}$$

where the Frobenius norm $||M||_F$ of a matrix $M = (M_{ij})$ is given by $||M||_F = \sqrt{\sum_{ij} M_{ij}^2}$. This problem has a closed-form solution through SVD [16]. For a general form of matrix approximation, Ye [1] proposed a *matrix space model* for low rank approximation of matrices called the generalized low rank approximation of matrices (GLRAM), which can be stated as follows.

Let $A_i \in \mathbb{R}^{n\times m}$, for $i = 1, 2, \cdots, N.$ be the N data points in the training set. We aim to compute two matrices $L \in \mathbb{R}^{n\times r}$ and $R \in \mathbb{R}^{m\times c}$ with orthonormal columns and N matrices $M_i \in \mathbb{R}^{r\times c}$ for $i = 1, 2, \cdots, N.$ such that LM_iR^T

approximates A_i well for all i. Mathematically, this can be formulated as an optimization problem

$$\min_{L^T L=I_r, R^T R=I_c} \sum_{i=1}^{N} ||A_i - LM_i R^\top||_F^2 \tag{2}$$

Actually, there is no closed form solution for the proposed GLRAM. In order to find an optimal solution for this GLRAM, Ye [1] has proved that the proposed GLRAM is equivalent to the following optimization problem.

$$\max_{L,R} J(L,R) = \sum_{i=1}^{N} ||L^\top A_i R||_F^2 = \sum_{i=1}^{N} Tr\,[\,L^\top A_i R R^\top A_i{}^\top L] \tag{3}$$

where L and R must satisfy $LL^\top = I_r$, $RR^\top = I_c$.

Instead of solving the above optimization problem directly, Ye [1] observed the following fact in Theorem 1 and proposed the **Algorithm GLRAM** stated below.

Theorem 1. *Let L, R and $\{M_i\}_{i=1}^N$ be the optimal solution to the minimization problem in GLRAM. Then*

(1) For a given R, L consists of the r eigenvectors of matrix

$$M_L(R) = \sum_{i=1}^{N} A_i R R^\top A_i^\top \tag{4}$$

corresponding to the largest r eigenvalues.

(2) For a given L, R consists of the c eigenvectors of the matrix

$$M_R(L) = \sum_{i=1}^{N} A_i^\top L L^\top A_i \tag{5}$$

corresponding to the largest c eigenvalues.

Based on these observations, the following algorithm for solving the GLRAM is proposed in Ye [1], which is called the **Algorithm GLRAM**.

Algorithm GLRAM
Input: matrices $\{A_i\}$, r and c.
Output: matrices L, R and $\{M_i\}$
1. Obtain initial L_0 for L and set $i = 1$.
2. While not convergent
3. From the matrix $M_R(L) = \sum_{j=1}^{N} A_j^\top L_{i-1} L_{i-1}^\top A_j$ compute the c eigenvectors $\{\phi_j^R\}_{j=1}^c$ of M_R corresponding to the largest c eigenvalues.
4. Let $R_i = [\phi_1^R, \cdots, \phi_c^R]$
5. From the matrix $M_L(R) = \sum_{j=1}^{N} A_j R_i R_i^\top A_j^\top$ compute the r eigenvectors $\{\phi_j^L\}_{j=1}^r$ corresponding to the largest r eigenvalues.

6. Let $L_i = [\phi_1^L, \cdots, \phi_r^L]$
7. i=i+1
8. Endwhile
9. $L = L_{i-1}$
10. $R = R_{i-1}$
11. $M_j = L^\top A_j R$

It should be noted here that the convergence of this algorithm can not be guaranteed theoretically in general and therefore the implementation from step 2 to step 8 may be never end. In order to investigate the convergence issue of the proposed algorithm, the following criteria is defined in Ye [1].

$$\mathrm{R}MSRE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} ||A_i - LM_iR^\top||_F^2} \tag{6}$$

where RMSRE stands for the *Root Mean Square Reconstruction Error*. It is easy to show that each iteration for the **Algorithm GLRAM** will decrease the RMSRE value and also RMSRE is bounded. Based on these facts, the following convergence result is obtained in Ye [1].

Theorem 2. ***Algorithm GLRAM****monotonically decreases the RMSRE value, hence it converges.*

Actually, Theorem 2 is not rigorously right in mathematical perspective though it is regarded as acceptable in engineering applications. To explain this issue clearly, we can rewrite RMSRE as $RMSRE(L, R)$ being a function of variables L and R. Then, one can see that the produced sequences $\{L_i, R_i\}$ will decrease the value $\{RMSRE(L_i, R_i)\}$ and this can ONLY guarantee that the sequence $\{RMSRE(L_i, R_i)\}$ will converge mathematically. In this case, one can not conclude the convergence of the sequence $\{L_i, R_i\}$ itself. However, if we assume that RMSRE(L,R) only has one unique optimal solution (L, R), Theorem 2 will be true in general.

In other words, the convergence analysis for the **Algorithm GLRAM** is not sufficient in Ye [1]. In order to further discuss the convergence issue, Lu [2] derived a necessary condition for the GLRAM problem. Lu [2] considered the optimization problem on manifolds due to the constraints for $L^\top L = I_r$ and $RR^\top = I_c$. In order to solve this constrained optimization problem, Lu [2] first derived the corresponding tangent spaces via using the manifold theory [17].

In order derive the tangent spaces for L and R, one can see that the constraints for L and R are made of Stiefel manifolds. Lu [2] obtained a necessary condition for the optimality of GLRAM problem is.

$$\begin{cases} (I - LL^\top)L^* = 0 \\ (I - RR^\top)R^* = 0 \end{cases} \tag{7}$$

We rewrite the necessary conditions as follows by substituting L^* and R^*, and by multiplying $L^\top$ and $R^\top$ on the right side of above equations, using the notations $M_R(L)$ and $M_L(R)$ as in Theorem 1, one can obtain the following.

$$\begin{cases} LL^\top M_L(R)LL^\top = M_L(R)LL^\top \\ RR^\top)M_R(L)RR^\top = M_R(L)RR^\top \end{cases} \tag{8}$$

With a given initial value L_0, if we denote the sequence produced by the **Algorithm GLRAM** as $\{L_{i-1}\}$ and $\{R_i\}$ for $i = 1, 2, \cdots$. Then, one can easily testify that they will satisfy the following equations

$$\begin{cases} L_{i+1}L_{i+1}^\top M_L(R_i)L_{i+1}L_{i+1}^\top = M_L(R_i)L_{i+1}L_{i+1}^\top \\ R_iR_i^\top)M_R(L_{i-1})R_iR_i^\top = M_R(L_{i-1})R_iR_i^\top \end{cases} \tag{9}$$

This implies that if the sequences $\{L_{i-1}\}$ and $\{R_i\}$ converge, their limits will satisfy the necessary conditions for the optimality of GLRAM. In addition, every two iterations in the **Algorithm GLRAM** will satisfy the necessary condition as shown in (9). However the convergence of these sequences can not be guaranteed theoretically in general due to their complex nonlinearity.

Mathematically speaking, L and R should be treated equally in (8) as two independent variables. In this case, one will derive sequences $\{L_i\}$ and $\{R_i\}$ by using the **Algorithm GLRAM** interchangeably .

3 2DPCA

In this section, we briefly outline the standard procedure of building the eigen vector space from a set of training images and then develop its incremental version. We represent input images as matrices.

Let $A_i \in \mathbb{R}^{m\times n}$, $i = 1 \ldots M$, where m, n are the number of pixels in the image, and M is the number of the images. We adopt the following criterion as in [3]:

$$J(X_r) = tr(S_{X_r}) = tr\{X_r^\top[\mathbf{E}(A - \mathbf{E}A)^\top(A - \mathbf{E}A)]X_r\} \tag{10}$$

where S_{X_r} is the covariance matrix of A_i (i=1,2...M) with the projection matrix X_r and $\mathbf{E}$ is the expectation of a stochastic variable. In fact, the covariance matrix $G \in \mathbb{R}^{n\times n}$ with M images is:

$$G = \frac{1}{M}\sum_{i=1}^{M}(A_j - \overline{A}_M)^\top(A_j - \overline{A}_M) \tag{11}$$

where $\overline{A}_M = \frac{1}{M}\sum_{i=1}^{M} A_i$ is the mean image matrix of the M training samples. Alternatively the criterion in (1) can be expressed by the following:

$$J(X_r) = tr(X_r^\top G X_r) \tag{12}$$

where X_r is a unitary column vector. The matrix X_r that maximizes the criterion is called the optimal projection axis. The optimal projection $X_{r_{opt}}$ is a set

of unitary vectors that maximizes $J(X)$. i.e the eigenvectors of G corresponding to the large eigenvalues [1]. We usually choose a subset of only d eigenvectors corresponding to larger eigenvalues to be included in the model, that is $\{X_{r_1} \dots X_{r_d}\} = \arg\max J(X_r)$ satisfying $X_{r_i}^\top X_{r_j} = 0(\ i \neq j,\ i, j = 1, 2, \dots, d)$. Each image can thus be optimally approximated in the least-squares sense up to a predefined reconstruction error. Every input image A_i will project into a point in the d-dimensional subspace spanned by the selected eigen matrix X_r [3].

4 Double Sides 2DPCA

Based on investigations from section 2 we could find two convergent matrices R and L through iteration algorithm of GLRAM. And the 2DPCA in section 3 is using SVD decomposition of covariance matrix $G_r = \frac{1}{M}\sum_{i=1}^{M}(A_j - \overline{A}_M)^\top(A_j - \overline{A}_M)$ to obtain a feature matrix R' similar to the R in GLRAM algorithm. We can ALSO get another feature matrix L' similaR TO THE L IN GLRAM using SVD decomposition of covariance matrix $G_l = \frac{1}{M}\sum_{i=1}^{M}(A_j - \overline{A}_M)(A_j - \overline{A}_M)^\top$. We then use a feature matrix R' and feature matrix L' to approximate R and L in GLRAM algorithm, respectively, we called this algorithm as doubleside 2DPCA algorithm. Experiments show that the performances of doubleside 2DPCA is as good as the GLRAM, and it save much time in recognition.

5 Doubleside Incremental 2DPCA

Let us now turn to the incremental version of the algorithm. Assume we have already build a eigen matrix $X_r = [X_{r_i}]$, $i = 1, 2, \dots, d$ with the input images A_i , $i = 1, \dots, M$. Actually X_r is the first d vectors of the following solution

$$G_r X_{r_w} = X_{r_w} \Lambda \tag{13}$$

where Λ is a diagonal matrix of eigenvalues with decreasing order. Incremental building of eigen matrices requires us to update X_r after taking a new input image A_{M+1} as input.

The new eigenspace model can be obtained incrementally by using Eigen Value Decomposition (EVD) after adding a new matrix A_{M+1} and is:

$$G'_r X'_{r_w} = X'_{r_w} \Lambda' \tag{14}$$

where X'_r can be chosen from X'_{r_w} as the first $d+1$ columns. Recall that we can update the mean with $\overline{A}_M$ and A_{M+1} as following.

$$\overline{A}_{M+1} = \frac{1}{M+1}(M\overline{A}_M + A_{M+1}) \tag{15}$$

So, we can update the covariance matrix G'_r from G_r, $A_{M+1} - \overline{A}_M$ as following.

$$G'_r = \frac{1}{M+1}\sum_{i=1}^{M+1}(A_j - \overline{A}_{M+1})^\top(A_j - \overline{A}_{M+1})$$

$$= \frac{M}{M+1}G + \frac{M^2}{(M+1)^3}(y)^\intercal y \tag{16}$$

where $y = A_{M+1} - \overline{A}_M$. From the above analysis, we can obtain X_r' after we get a new image A_{M+1} with the previous information G_r and $\overline{A}_M$ by solving equation (14). One important question is: whether it is necessary to add one more dimension after adding the image A_{M+1}. Since the updated X_r' may bring no useful information or even lead to worse performance if A_{M+1} could not bring effective information compared to previous A_i, $i = 1, 2, \cdots, M$.

Based on above analysis, we need to update the eigen matrix after adding a new matrix A_{M+1} based on the actual information brought by A_{M+1}. In order to do this, we first compute an orthogonal residual vector h_{M+1} based on the effective information y in (16); The new observation A_{M+1} is projected into the eigenspace model by X_r to give a d-dimensional matrix g. g can be reconstructed into $m \times n$ dimensional space, but with the loss represented by the residue matrix h_{M+1}.

$$g = X_r^\top (A_{M+1} - \overline{A}_M) \tag{17}$$

$$h_{M+1} = (A_{M+1} - \overline{A}_M) - Xg \tag{18}$$

and we normalize (18) to obtain the following.

$$\hat{h}_{M+1} = \begin{cases} \frac{h_{M+1}}{||h_{M+1}||_F} & ||h_{M+1}||_F > \delta \\ 0 & otherwise \end{cases} \tag{19}$$

Now this reconstruction error is used to determine whether we need to update X_r. If $\hat{h}_{M+1}$ is very small and negligible, it indicates that the added image A_{M+1} did not bring any effective information and we do not need to update X_r. Otherwise, we need to update X_r based on this effective information $\hat{h}_{M+1}$. Recall that $\hat{h}_{M+1}$ is a matrix and we can add only one more dimension for X_r. In order to represent the effective information with a vector, we construct a column vector h using the mean vector of columns in $\hat{h}_{M+1}$ as h.

In this case, the vector h is a representation for the effective information brought by A_{M+1}. Next, we will use this effective information to update the eigen matrix. Since h may not orthogonal to X_r, as did in [1], we update X_r in the following way,

$$X_r^{'} = [X_r, h]R \tag{20}$$

where $R \in R^{m \times (d+1)}$ is rotation matrix. In order to determine R optimally, we need $X_r^{'}$ to be the first $d+1$ eigenvectors of $G^{'}$ and this gives

$$[X_r, h]^\intercal (\frac{M}{M+1}G_r + \frac{M^2}{(M+1)^3}(y)^\intercal y)[X_r, h]R = R\Lambda^{'} \tag{21}$$

after substitution of Equations (16) and (20) into (14). This is a $(d+1) \times (d+1)$ eigen problem with solution eigenvector R and eigenvalues $\Lambda^{'} \in \mathbb{R}^{(d+1)\times(d+1)}$.

Note that we have reduced dimensions from n to $(d+1)$ discarding the eigenvalues in $\Lambda_w^{'}$ already deemed negligible.

The left hand side comprises two additive terms. Up to a scale factor, the first of those terms is:

$$[X_r, h]^\intercal G_r [X_r, h] = \begin{bmatrix} X_r^\intercal G_r X_r & X_r^\intercal G_r h \\ h^\intercal G_r X_r & h^\intercal G_r h \end{bmatrix} \approx \begin{bmatrix} \Lambda & 0 \\ 0^\intercal & 0 \end{bmatrix} \tag{22}$$

where 0 is a d-dimension vector of zeros. In [1] it is proved that $G_r \hat{h} \approx 0$, where $\hat{h}$ is normalized residue vector. In our paper, $\hat{h}$ is a residue matrix, we use the mean of columns in $\hat{h}$ as h, so $G_r h \approx 0$. Therefore, $X^\intercal G_r h \approx 0$, $h^\intercal G_r X \approx 0$, $h^\intercal G_r h \approx 0$.

Similarly, The second term will be

$$[X_r, h]^\intercal ((y)^\intercal y)[X_r, h] = \begin{bmatrix} \alpha^\intercal \alpha & \alpha^\intercal \gamma \\ \gamma^\intercal \alpha & \gamma^\intercal \gamma \end{bmatrix} \tag{23}$$

where $\gamma = (A_{M+1} - \overline{A}_M)h$ and $\alpha = (A_{M+1} - \overline{A}_M)X_r$

Thus, $R \in \mathbb{R}^{(d+1)\times(d+1)}$ of (19) will be the solution of the following form

$$DR = R\Lambda^{'} \tag{24}$$

where $D \in \mathbb{R}^{(d+1)\times(d+1)}$ is as following.

$$D = \frac{M}{M+1} \begin{bmatrix} \Lambda & 0 \\ 0^\top & 0 \end{bmatrix} + \frac{M^2}{(M+1)^3} \begin{bmatrix} \alpha^\intercal \alpha & \alpha^\intercal \gamma \\ \gamma^\intercal \alpha & \gamma^\intercal \gamma \end{bmatrix} \tag{25}$$

Concludingly, we can obtain an updated $X_r^{'}$ once we solved (25) for R. In the same way, we can obtain an updated $X_l^{'}$. Then, we can obtain an updated $Y^{'}$ using $X_r^{'}$ and $X_l^{'}$ feature matrices.

6 Proposed Incremental Face Recognition Algorithm

In this section, we will propose an updated face recognition algorithm based on the incremental projection derived in last section.

In the eigenspace method, 2DPCA can be used to reconstruct the images [3] as following. The M images A_i $i = 1, \ldots, M$, have produced a dr dimensional eigen matrix $[X_{r_1}, X_{r_2}, \ldots, X_{r_{dr}}]$ with the image covariance matrix G_r and dl dimensional eigen matrix $[X_{l_1}, X_{l_2} \ldots, X_{l_{dl}}]$ with the image covariance matrix G_r and G_l After the image samples are projected onto these axes, the resulting principal component matrices are $Y_k = X_l A_k X_r$ $(k = 1 \ldots, M)$, and can be approximately reconstructed by $\tilde{A}_k = X_l^\top Y_k X_r^\top$ $(k = 1 \ldots, M)$ $(dr \leq n, dl \leq m)$. The images are represented with coefficient matrix $\tilde{A}_k$ and can be approximated by $\hat{A}_k = \tilde{A}_{k(M)} + \bar{A}_M$ $(k = 1 \ldots, M)$. The subscript (M) denotes the discrete time line of adding the images.

When a new observation A_{M+1} arrives, we compute the new mean using (15) and the new covariance matrix using (16). Further, we can construct the

intermediate matrix D (25), and solve the eigenproblem (24), respectively. This produces a new subspace of eigenmatrices $X_r^{'}$ and $X_l^{'}$.

In order to remap the coefficients $\hat{A}_{i(M)}$ into this new subspace, we first compute an auxiliary matrix η .

$$\eta = X_l^{'}((\bar{A}_M - \bar{A}_{M+1})X_r^{'}) \tag{26}$$

which is then used in the computation of all coefficients,

$$\hat{Y}_{i_{M+1}} = R_l[\begin{bmatrix} Y_{i_M} & 0 \\ 0 & 0 \end{bmatrix} R_r], i = 1, \ldots, M+1 \tag{27}$$

The above transformation produces a representation with $(dl+1) \times (dr+1)$ dimension. In this case, the approximations of previously images $\hat{A}_k$ $(k = 1 \ldots, M)$ and the new observation A_{M+1} can be fully reconstructed from this representation as.

$$\hat{A}_{i_{M+1}} = X_l^{'\prime}\hat{Y}_{i_{M+1}}X_r^{\prime} + \overline{A}_{M+1} \tag{28}$$

In next section, we will do some experiments with different databases to show the effectiveness of the proposed update algorithm.

Based on the above analysis, we outline our incremental algorithm for face recognition as the follows:

Algorithm 1

1: Give M training images $A_i(i = 1....M)$, calculate the mean $\bar{A}_M$, covariance matrices G_l, G_r and eigenspace X_l, X_r
2: Adding a image A_{M+1}, updating the mean $\bar{A}_{M+1}$ with $\bar{A}_M$ and A_{M+1}.
3: Calculate the vector h_r, if $h_r > \delta_r$, then $X_r^{'} = [X, h_r]R_r$, else $X_r^{'} = X_r$. Calculate the vector h_l, if $h_l > \delta_l$, then $X_l^{'} = [X, h_l]R_l$, else $X_l^{'} = X_l$.
4: goto 2.

Concludingly, we have achieved the updated algorithm.

Algorithm 2

1) We have updated eigenspace matrices

$$X_l^{'} = [X, h_l]R_l \; and \; X_r^{'} = [X, h_r]R_r$$

where R_r and R_l are obtained from (24), respectively, if the effective information is significant.

2) We have updated projection for training samples

$$\hat{Y}_{i_{M+1}} = R_l[[Y_{i_M}, 0]\, R_r, 0] + \eta.$$

where η is given in (26), if update is necessary.

3) In Algorithm 1, the average value of training samples must be updated even the eigenspace matrix is not changed. This is to make sure that the projections of training samples are updated properly in (27), in the case of multi-step updations.

7 The Experimentation

In this section, we will carry out several experiments to demonstrate the performance of the incremental 2DPCA face recognition algorithms. As to the databases, we used two types of input images. One is the AMP face database, which contains 975 images of 13 individuals(each person has 75 different images) under various facial expressions and lighting conditions with each images being cropped and resized to 64×64 pixels in this experiment. The second one is ORL face database, which includes 400 images of 40 individuals(each person has 10 different images) under various facial expressions and lighting conditions with each images being cropped and resized to 112×92 pixels in this experiment. The third one is yale face database, which includes 165 images of 15 individuals(each person has 11 different images) under various facial expressions and lighting conditions with each images being cropped and resized to 231×195 pixels in this experiment. The other database is the YaleB face database [14], which contains 650 face images with dimension 480×640 of 10 people , including frontal views of face with different facial expressions, lighting conditions with background. With these databases, we will conduct four experiments. The first one is to compare the result of face recognition with the batch algorithm, GLRAM algorithm and double sides 2DPCA. The second experiment is to demonstrate the performance of the RMSRE, it stands for the *Root Mean Square Reconstruction Error* for face recognition between GLRAM algorithm and double sides 2DPCA. The third one is to compare the computation cost in the procedure of recognition among the batch algorithm, GLRAM algorithm and double sides 2DPCA. The final experiment is to show computation cost and demonstrate the efficiency of the updated algorithm.

Top five images of each individuals in all the experiments were used as trainings and the others regard as testings. All testings were properly recognized when the feature d is more than 3 in the AMP database; In Yale database, the recognition accuracy is always 85.6 when the feature d is more than 3. The other two database experiment results show in the following sections.

7.1 Performance Comparisons of Face Recognition Algorithms

We select $d = 3, 4, 5, 6, 7, 8$ and $d = 40, 41, 42, 43, 44, 45$ in the database ORL and YaleB, respectively, since all algorithms reach high recognition accuracy in those areas. The rate of recognition accuracy in ORL database is show in Table 1 and the YaleB database is in Table 2.

Table 1. Recognition accuracy % in ORL database $d = 3, 4, ..., 7, 8$

algorithm	3	4	5	6	7	8
GLRAM	73	85	84.5	86	85.5	85
Batch	80.5	81	80	83	83	82.5
Doubleside	74.5	81	82	81.5	86	84.5

Table 2. Recognition accuracy % in YaleB database $d = 40, 41, ..., 45$

algorithm	40	41	42	43	44	45
GLRAM	74.7	74.9	74.9	74.9	74.9	74.9
Batch	74.9	74.9	74.9	74.9	74.9	74.9
Doubleside	74.8	74.9	74.9	74.9	74.9	74.9

It is shown that performance in GLRAM and doubleside 2DPCA are a little bit of higher than 2DPCA algorithm in ORL face database and the performance will drop after d=8. The performance in Yale B database are same when d arrives at 41, and when d reaches 60 the performance of them is 95.1.

7.2 Performance of the RMSRE

It is easy to know that the RMSRE values (stands for the *Root Mean Square Reconstruction Error)* are very closely, the RMSRE value will decrease as d increase and also RMSRE is bounded.

Table 3. The RMSRE in ORL database $d = 3, 4, ..., 8$

algorithm	3	4	5	6	7	8
GLRAM	39.7	37.7	35.4	33.9	32.1	30.7
Doubleside 2DPCA	39.9	38.1	35.5	34.0	32.1	30.8

Table 4. The RMSRE in YaleB database $d = 40, 41, ..., 45$

algorithm	40	41	42	43	44	45
GLRAM	42.4	41.8	41.5	40.9	40.3	39.8
Doubleside 2DPCA	42.4	42.0	41.6	41.1	40.7	40.2

7.3 Comparisons of Computational Costs

It is easy to show that the computation costs in recognition for 2DPCA is doupled than the doublesides 2DPCA and GLRAM in the ORL face database, and it is even more costive in YaleB database due to the fact that 2DPCA only reduces dimensions in one side and doubleside 2DPCA will reduce dimensions on both sides.

7.4 Comparisons of the Batched 2DPCA, Incremental 2DPCA, Doubleside 2DPCA(DS2DPCA), Doubleside Incremental 2DPCA(DSI2DPCA) and GLRAM

In this section, the YaleB face databases are used for experiments to compare the recognition performance of Batch 2DPCA, Incremental 2DPCA, Doubleside

Table 5. Computation Cost in ORL database $d = 3, 4, ..., 8$

algorithm	3	4	5	6	7	8
GLRAM	3.58	3.97	4.78	5.48	5.85	6.60
2DPCA	5.42	6.44	7.87	9.06	11.47	11.68
Doubleside 2DPCA	3.67	4.02	5.20	5.29	5.97	6.94

Table 6. Computational Costs in YaleB database $d = 40, 41, ..., 45$

algorithm	40	41	42	43	44	45
GLRAM	102	106	117	119	139	155
2DPCA	318	334	342	391	379	400
Doubleside 2DPCA	97	110	120	121	131	137

2DPCA and Doubleside Incremental 2DPCA. In all the experiments for YaleB database, we set d=41, 42, 43, 44, 45 and use the first 8 images of each individual as training. The ninth image of each person is used for incoming images and all the images after tenth image are used for testing in YaleB database. We use 60 images per person for testing and use 50 images per individual for testing in YaleB database. The batch algorithm(2DPCA and DS2DPCA) performs a single SVD on the covariance matrix containing all images in ensemble and obtain an eigen matrix. Hence, it represents the best possible scenario for 2DPCA in terms of recognition and reconstruction. So we need to compare its recognition performance with that produced by the incemental 2DPCA(I2DPCA and DSI2DPCA).

The YaleB face images database is a typical face database since it includes many faces with different expressions, light illuminating conditions etc. Also the accuracy of face recognition methods using this database is usually very low [13]. The experiments of continued adding the ninth image A9 to E9 of the top of five persons were done, we obtain the results which are the number of total correct matched in ten groups divided by the number of all testings with each d using the four techniques. The results are shown in Fig.1.

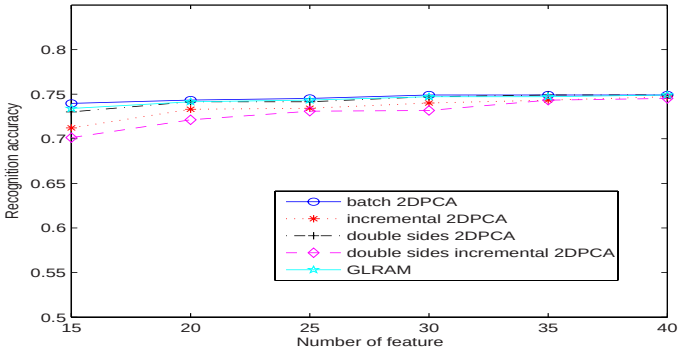


Fig. 1. Comparision of five approaches under the YaleB database(d=[15 20 ... 40])

7.5 Computational Cost Comparisons

In this section, we compare the computational time used to recognize all testing via using different techniques in the YaleB database. Here we use a PC machine (CPU: PIV 1.3GHz, RAM: 256M) to run Matlab and set set $d = 41, 42, 43, 44, 45$ and use initial training of 80 images 1 to 8 each person and we recognized the testing of 540 images from 10 to 64 each individuals, respectively. After adding the ninth image A9 to E9 of the top of five persons in turn, we measure the times to recognize the testing using different techniques as listed in the table 7. The time listed is the average time we run each program ten times (time unit : second).

Table 7. Comparison of CPU time in YaleB database $d = 41, 42, ..., 45$

algorithm	A9	B9	C9	D9	E9
batch	3752	3791	3799	3860	3668
I2DPCA	3751	3788	3799	3858	3664
DS2DPCA	147.8	152.1	153.4	158.5	156.1
DSI2DPCA	148.5	151.6	152.7	156.2	157.4
GLRAM	163.7	167.9	169.8	173.7	181.4

It can be seen that the incremental algorithm is about much faster than the batch algorithm for YaleB database. The GLRAM algorithm is a slightly slower than the incremental algorithm, maybe because there are some memory needs for convergent matrices, but it needs much more time for solutiong of convergent matrices.

8 Conclusions

In this paper, we developed an updated two sides 2DPCA algorithm and demonstrated its effectiveness in face recognition. The complexity of the algorithm is analyzed and the performance is compared among the batch algorithm, incremental 2DPCA, doubleside 2DPCA and GLRAM algorithm. The results showed that the proposed algorithm is working well for face recognition. We believe they are applicable in many applications when incremental learning is necessary.

In the future, we will investigate the convergence of these algorithms and apply these techniques to other surveillance applications.

References

[1] Ye, J.P.: Generalized Low Rank Approximations of Matrices. In: The Proceedings of Twenty-first International Conference on Machine Learning, Banff, Alberta, Canada, pp. 887–894 (2004)

[2] Lu, C., Liu, W.Q., An, S.J.: Revisit to the Problem of Generalized Low Rank Approximation of Matrices. In: International Conference on Intelligent Computing. ICIC, pp. 450–460 (2006)

[3] Hall, P.M., Marshall, D., Martin, R.R.: Incremental Eigenanalysis for Classification. PAMA 22(9), 1042–1048 (2000)
[4] Artač, M., et al.: Incremental PCA for on-line visual learning and recognition. In: IEEE, 16th international conference on Pattern Recognition, pp. 781–784 (2002)
[5] Yang, J., et al.: Two-Dimentional PCA: a new approach of 2DPCA to appearance-based face representation and recognition. IEEE Tran. Pattern Analysisand Machine Intelligence 26(11), 131–137 (2004)
[6] Sirovich, L., Kirby, M.: Low-Dimensional Procedure for Characterization of HumanFaces. J. Optical Soc. Am. 4, 519–524 (1987)
[7] Kirby, M., Sirovich, L.: Application of the KL Procedure for the Characterization of Human Faces. IEEE Trans. Pattern Analysisand Machine Intelligence 12, 103–108 (1990)
[8] Grudin, M.A.: On Internal Representations in Face Recognition Systems. Pattern Recognition 33, 1161–1177 (2000)
[9] Zhao, L., Yang, Y.: Theoretical Analysis of Illumination in PCA-Based Vision Systems. Pattern Recognition 32, 547–564 (1999)
[10] Valentin, D., Abdi, H., O'Toole, A.J., Cottrell, G.W.: Connectionist Models of Face Processing: a Survey. Pattern Recognition 27, 1209–1230 (1994)
[11] Hall, P., Marshall, D., Martin, R.: Merging and splitting eigenspace models. PAMI 22, 1042–1048 (2000)
[12] Winkeler, J., Manjunath, B.S., Chandrasekaran, S.: Subset selection for active object recognition. In: CVPR, June 2, pp. 511–516. IEEE Computer Society Press, Los Alamitos (1999)
[13] Tjahyadi, R., Liu, W.Q., Venkatesh, S.: Automatic Parameter selection for Eigenface. In: Proceeding of 6th International Conference on Optimization Techniques and Applications. ICOTA (2004)
[14] `ftp://plucky.cs.yale.edu/CVC/pub/images/yalefaceB/Tarsets`
[15] `http://www.library.cornell.edu/nr/bookcpdf.html`
[16] Golub, G.H., Van Loan, C.F.: Matrix computations, 3rd edn. The Johns Hopkins University Press, Baltimore (1996)
[17] Helmke, U., Moore, J.B.: Optimization and Dynamic Systems. Springer, London (1994)
[18] Lu, C., Liu, W.Q., An, S.J., Venkatesh, S.: Face Recognition via Incremental 2DPCA. In: International Joint Conference on AI. IJCAI, pp. 19–24 (2007)

Ear Recognition with Variant Poses Using Locally Linear Embedding

Zhaoxia Xie and Zhichun Mu

School of Information Engineering, University of Science and Technology Beijing, 100083, China
xiezhaox@163.com, mu@ies.ustb.edu.cn

Abstract. Ear recognition with variant poses is an important problem. In this paper locally linear embedding (LLE) is introduced considering the advantages of LLE, by analyzing the shortcomings of most ear recognition methods currently when dealing with pose variations. Experimental results demonstrate that applying LLE for ear recognition with variant poses is feasible and can obtain the better recognition rate, which shows the validity of this algorithm.

Keywords: Ear recognition; locally linear embedding; pose variations.

1 Introduction

In recently years, ear recognition technology has been developed from the feasible research to the stage of how to enhance ear recognition performance further, such as 3D ear recognition [1], ear recognition with occlusion [2], and ear recognition with variant poses etc. This paper addresses the issue of ear recognition with variant poses using 2D ear images without occlusion under invariant illumination condition.

The most known ear recognition was made by Alfred Iannarelli [3]. Moreno et al. [4] used three neural net approaches for ear recognition. Burge and Burger [5] have researched ear biometrics with adjacency graph built from Voronoi diagram of its curve segments. Hurley et al. [6] have provided force field transformations for ear recognition. Mu et al. [7] presented a long axis based on shape and structural feature extraction method (LABSSFEM). All these methods are called the method based on geometric feature, which needs to extract the geometric feature of ear images and is obviously influenced by pose variations, hence not suitable for multi-pose ear recognition. Along the different line, PCA [8] and KPCA [9] which are called the method based on algebraic feature are used for ear recognition. However, when data points are distributed in a nonlinear way such as pose variations, PCA will fail to discover the nonlinearity hidden in data points. Although KPCA is the method based on nonlinear technique, but projection results obtained by KPCA aren't visual. As the nonlinear algebraic approach, locally linear embedding (LLE) algorithm is proposed [10]. The implicit idea behind LLE is to project input points in high dimension space into a single global coordinate system of lower dimension, preserving neighboring relationships and discovering the underlying structure of input points.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 460–465, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The remainder of this paper is organized as follows: Section 2 presents LLE algorithm. Experimental results are given in section 3. Finally, the conclusion and future work are discussed.

2 Locally Linear Embedding

2.1 Algorithm

Given the real-valued input vectors $\{X_1, X_2, \cdots, X_N\}$, $i = 1, \cdots, N$, each of dimensionality D ($X_i \in R^D$), sampled from a smooth underlying manifold. Output vectors $\{Y_1, Y_2, \cdots, Y_N\}$, $i = 1, \cdots, N$, will be obtained, each of dimensionality d ($Y_i \in R^d$), where $d << D$. LLE algorithm consists of three steps [10] and [11]:

Step 1: Select k neighbors of each input vector $\overrightarrow{X_i}$. In the simplest formulation of LLE, one identifies k nearest neighbors of every data point, as measured by Euclidean distance.

Step 2: Compute the weights W_{ij} that best reconstruct each data point $\overrightarrow{X_i}$ from its neighbors. Reconstruction errors are measured by the cost function in equation (1), which adds up the squared distance between all the data points and their reconstruction. To compute the weights W_{ij} , one can minimize the cost in equation (1) subject to two constraints: a sparseness constraint and an invariance constraint. The sparseness constraint is that each data point $\overrightarrow{X_i}$ is reconstructed only from its neighbors, enforcing W_{ij} =0 if $\overrightarrow{X_j}$ does not belong to the set of neighbors of $\overrightarrow{X_i}$ .The invariance constraint is that the rows of the weight matrix sum to one $\sum_j W_{ij} = 1$ .The optimal weights W_{ij} subject to these constraints are found by solving the constrained least squares problems.

$$E(W) = \sum_i \left| \overrightarrow{X_i} - \sum_j W_{ij} \overrightarrow{X_j} \right|^2 \tag{1}$$

Step3: Compute the vectors $\overrightarrow{Y_i}$ best reconstructed by the weights W_{ij} . In this step, each high dimensional input point $\overrightarrow{X_i}$ is mapped to a low dimensional output $\overrightarrow{Y_i}$ representing global internal coordinates on the manifold. This is done by choosing the m-dimensional coordinates of each output point $\overrightarrow{Y_i}$ to minimize the embedding cost function in equation (2).

$$\Phi(Y) = \sum_i \left| \overrightarrow{Y_i} - \sum_j W_{ij} \overrightarrow{Y_j} \right|^2 \tag{2}$$

Fig.1 gives the example of using LLE on the synthetic data set. On the S-Curve surface in R^3, 2000 sample points are generated randomly. Lower-dimension embedding result are obtained with k =37 using LLE.

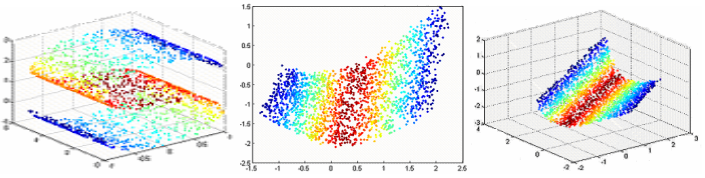


Fig. 1. 2000 sample points are generated randomly (left); two-dimensional embedding result (middle); three-dimensional embedding result (right)

2.2 Examples of Ear Feature Extraction

Fig. 2 shows the embedding results using LLE. Ear images are selected from USTB ear database, which will describe in the section3.Here, ear angle varies from 5 degree to 40 degree every other 5 under the left-turning pose, namely there are totally 79*8 ear images used in this experiment.

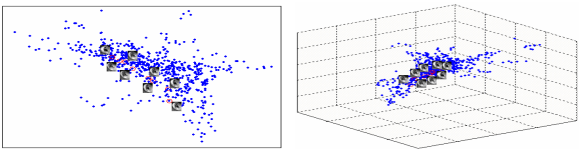


Fig. 2. Embedding results obtained in the two dimensional space using ear images (left); embedding results obtained in the three dimensional space using ear images (right)

3 Experiments and Results

3.1 Ear Data Set

Experiments are conducted on the USTB ear database which has been developed in our lab. Images containing head in the white background are acquired by a color CCD camera fixed on the tripod. The individuals were positioned seated at approximately 150cm from the lens plane, shooting head from side view. Nineteen images corresponding to 19 different poses are captured for individual. Images of each individual without occlusion are taken every other 5° in the range of 45° under the invariant illumination condition including left turn and right turn. This database now contains the images of 79 individuals, namely 1501images totally. The size of original 24bit color image is 768*576 pixels. In experiments, firstly the color images are transformed into gray images, then using the long axis of outer ear contour [9] to manually segment and rotate ear from original images. After wiener filtering and linear interpolation, the size of ear images is normalized to 45*24 pixels. Fig. 3 gives the example of this database and the example of ear images normalized.

Fig. 3. The example of one individual with 19 poses in the USTB ear database (left); the illustration of ear images normalized (right)

3.2 Experiments Based on Scheme One

The first scheme is designed as follows: the frontal ear image of every individual, namely 79 ear images consist of the training set. Ear images with pose variations are used for testing samples respectively, i.e. there are 18 test set, each of 79 ear images. All experiments are based on the NN classifier. The value of k is set to 12.

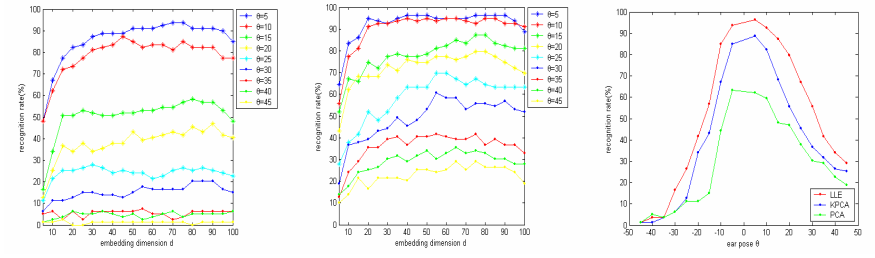


Fig. 4. The recognition rate under left-turning pose (left); the recognition rate under right-turning pose (middle); the recognition rate obtained using LLE, PCA and KPCA respectively when the embedding dimension is 75(right)

Fig.4 (left) and (middle) give experiment results, where the horizontal axis represents embedding dimension and the vertical axis represents the recognition rate. Firstly, the recognition rate decreases when the angel increases. Secondly, the recognition rate under right-turning pose is generally better than the recognition rate under left-turning pose at the same angle. Thirdly, the recognition rate that achieves to 80% is defined as the valid recognition rate. In this case the valid recognition range is from -10° to $+15^\circ$; '-'denotes left turn, and '+'denotes right turn. Based on the first scheme and the same experiment data, when the embedding dimension is 75, recognition results are given in Fig. 4(right) using LLE, PCA and KPCA respectively. The horizontal axis represents ear pose and the vertical axis represents recognition rate; '-'denotes left turn. Compared to KPCA and PCA, LLE has obvious advantages for ear recognition with poses changing. It shows that the better recognition rate is obtained at every angle. The reason is that LLE can preserve the local geometrical property of nonlinear data, and has advantages of dealing with nonlinear data.

3.3 Experiments Based on Scheme Two

The second scheme is given as follows: the frontal ear image, ear image with -5° and ear image with $+5^\circ$ of every individual ('-'denotes left turn, and '+'denotes right turn),

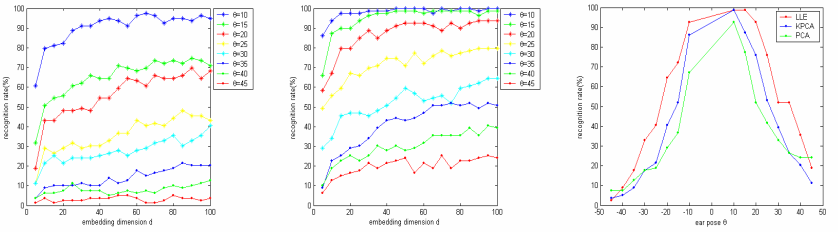


Fig. 5. The recognition rate under left-turning pose (left); the recognition rate under right-turning pose (middle); the recognition rate obtained using LLE, PCA and KPCA respectively when the embedding dimension is 75(right)

namely 79*3 ear images consist of the training set. Ear images with other poses are used for testing samples respectively, i.e. there are 16 test set, each of 79 ear images. All experiments are based on the NN classifier. The value of k is set to 12.

Fig.5 (left) and (middle) give experiment results, where the horizontal axis represents the embedding dimension and the vertical axis represents the recognition rate. From Fig.5 one can see that recognition rate decreases when the angel increases; and the information of ear images under right-turning pose is more abundant than that of ear images under left-turning pose. Based on the second scheme and the same experiment data, when the embedding dimension is 75, recognition results are given in Fig. 5(right) using PCA, KPCA and LLE respectively. The horizontal axis represents ear pose and the vertical axis represents the recognition rate; '-'denotes left turn. Compared to KPCA and PCA, the best recognition rate is again obtained using LLE, which demonstrates that LLE has advantages for ear recognition with variant poses.

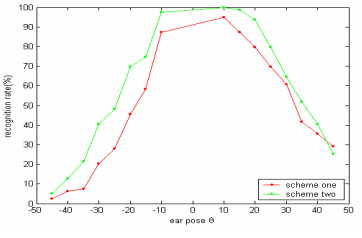


Fig. 6. Recognition results using two experimental schemes; the recognition rate denotes the highest recognition when the embedding dimension varies from 5 to 100 at every pose every other 5 dimensions

Fig. 6 gives the comparison of experimental results based on these two schemes. By contrast one can see that the recognition rate using scheme two is generally better than the recognition rate using scheme one at the same angle; here the recognition rate denotes the highest recognition when the embedding dimension varies from 5 to 100 every other 5 dimensions as for every pose. It shows that the recognition performance will enhance when the number of training set increases. It is impossible to get enough samples for every individual, yet acquiring one or three images of every individual is

feasible in real application. Experimental results demonstrate that the validity of LLE for ear recognition with variant poses.

4 Conclusion and Future Work

LLE algorithm which is suitable for ear recognition with variant poses is proposed. In order to demonstrate the validity of this algorithm, two different experiment schemes are designed. Experiment results show that one can obtain a satisfying recognition rate using LLE due to the effective ear feature extracted; hence the robustness of ear recognition is improved. However, neighbors of every input vector are selected according to Euclidean distance using LLE which is not suitable in the high dimensional space. Therefore, the next work will consider one reasonable similarity rule used for the choice of the neighbors of every data points.

Acknowledgements. This work is supported by: (1) the National Natural Science Foundation of China under the Grant No. 60375002, and No. 60573058; (2) the key discipline development program of Beijing Municipal Commission of Education No. XK100080537.

References

1. Chen, H., Bhanu, B.: Human Ear Recognition in 3D. IEEE Trans. Pattern Analysis and Machine Intelligence 29, 718–737 (2007)
2. Yuan, L., Mu, Z.C., Zhang, Y., Liu, K.: Ear Recognition Using Improved Non-negative Matrix Factorization. In: Proc of the 18th International Conference on Pattern Recognition, HongKong, China, vol. 4, pp. 501–504 (2006)
3. Iannarelli, A.: Ear identification. Paramount Publishing Company, USA (1989)
4. Moreno, B., Aánchez, A., Vélez, J.F.: On the Use of Outer Ear Images for Personal Identification in Security Applications. In: Proc. of IEEE 33rd Annual International Carnahan Conference on Security Technology, Madrid, Spain, pp. 469–476 (1999)
5. Burge, M., Burgem, W.: Ear Biometrics in Computer Vision. In: Proc. of the 15th International Conference of Pattern Recognition, Barcelona, Spain, pp. 822–826 (2000)
6. Hurleym, D.J., Nixon, M.S., Carter, J.N.: Force Field Feature Extraction for Ear Biometrics. Computer Vision and Image Understanding 98, 491–512 (2005)
7. Mu, Z.C., Yuan, L., Xu, Z.G., Xi, D.C., Qi, S.: Shape and Structural Feature Based Ear Recognition. In: Li, S.Z., Lai, J.-H., Tan, T., Feng, G.-C., Wang, Y. (eds.) SINOBIOMETRICS 2004. LNCS, vol. 3338, pp. 663–670. Springer, Heidelberg (2004)
8. Ping, Y., Bowyer, K.: Empirical Evaluation of Advanced Ear Biometrics. In: Proc of Empricial Evaluation Methods in Computer Vision, San Diego, pp. 56–59 (2005)
9. Yuan, L.: Study on Some Key Issues in Ear Recognition. University of Science and Technology Beijing (2006)
10. Roweis, S.T., Saul, L.K.: Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science 290, 2323–2326 (2000)
11. Saul, L.K., Roweis, S.T.: Think Globally, Fit locally: unsupervised Learning of Low Dimensional Manifolds. J. Machine Learning Research 4, 119–155 (2003)

Motion Detection with Background Clutter Suppression Based on KDE Model

Han Zhou, Zhi-yuan Zeng, and Jian-zhong Zhou

Digital Engineering and Simulation Research Center
Huazhong University of Science and Technology
Wuhan 430074, China
batigolgol@gmail.com

Abstract. Background modeling is an important component of visual surveillance system. In some complicated outdoor system, such as traffic scene in night, solutions to problems as illumination and shadow disturbance are provided. The kernel density estimation is exploited to estimate the probability density function of background intensity and then to classify the pixel into background or foreground scene. Toward the modeling of dynamic characteristics, a normalized color space is proposed as part of a five-dimensional feature space. And experiment demonstrates the performance of the proposed approach.

Keywords: Background modeling, Kernel density estimation, Motion detection, Normalized color space.

1 Introduction

Motion detection is the first step for tracking, target classification or some other visual surveillance applications. Background subtraction is a typical method for motion detection in video image sequences (taken from a static sensor) analysis , which is employed in the module of vehicle flow detection by comparing the current frame with the established background scene .The certain pixel intensity over time can be modeled by a zero mean Normal distribution $N(0,\sigma^2)$.

However, the modeling method such as Gaussian Mixture Model and Kalman Filter can only adapt to the slowly changing of background (like illumination changes). In many visual surveillance applications that work with outdoor scenes, the background of the scene contains many clutters such as non-static small objects.

C.R.Wren[1] proposed a single Gaussian to model the background pixel. Friedman[2] uses a mixture of three Normal distributions to model the visual properties in surveillance system. Three distributions are considered as road, shadow and vehicles.The idea in [3] goes further which uses multiple Gaussian to model the scene and develop a fast approximate method for updating the parameters of the background model incrementally. Such approaches are capable of dealing with multiple hypothesis for the background and can be useful in scenes which contain waving trees and ocean wave.

For the case of traffic circumstances in the night scene, background modeling is one of the key problems because the illumination from the lamp on the vehicle and

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 466–473, 2008.
© Springer-Verlag Berlin Heidelberg 2008

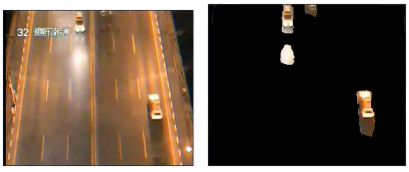


Fig. 1. The left one is the 13th frame from the original image sequences, the right one is the detection result with single Gaussian Model to model the background

shadow will affect the background scene, some pixels in illumination region will be detected as the moving object. As is shown below, the illumination from the left car is classified as foreground, and so is the shadow of the right car.

Regarding the background with clutter and illumination disturbance, we need to model a robust background, considering Kernel Density Estimator to model the background for the image sequences and it also can estimate the probability of observing pixel intensity values based on a sample of intensity values for each pixel.

A.Mittal[4] utilized a five-dimensional space, two components for the optical flow and three for the intensity in the normalized color space. Clutters like leaves moving and ocean's wave in the background scene are suppressed. And it defines the multivariate fixed bandwidth kernel estimator as:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\|H\|^{1/2}} K(H^{-1/2}(x - x_i)) \tag{1}$$

Yaser Sheikh[5] employed kernel density estimation as the first stage for subsequent MRF-MAP estimation based on Bayesian modeling. It is also defined in 5-dimensional space, but two components for the (x,y) , and the other three components for the RGB color space.

$$P(x \mid \varphi_b) = n^{-1}\sum_{i=1}^{n} \varphi_H(x - y_i) \tag{2}$$

And Gaussian kernel density is utilized for the KDE above in [5], so:

$$\varphi_H^{(N)}(x) = \mid H \mid^{-1/2} (2\pi)^{-d/2} \exp(-\frac{1}{2} x^T H^{-1} x) \tag{3}$$

Approach in [4] utilizes optical flow as two of the components in 5-dimensional feature space, which is computationally expensive and difficult to meet the need for real-time surveillance. Approach in [5] shows a good performance in background modeling and real-time quality, however, this approach doesn't mention the ability for illumination and shadow suppression.

In this paper, we propose an approach to build a robust background scene based on kernel density estimation, motion detection based on non-parametric KDE model is more appropriate when the sample points have variable uncertainties associated with them. And a multiple dimensional feature space which consists of intensity and motion feature, is utilized to represent a certain pixel.

This paper is organized as follows. Section 2 describes kernel density estimation with variable bandwidth, which utilizes both sample points and observation points. In section 3, the classification mechanism for the observed data based on the estimation probability in section 2 is proposed. Such a mechanism decides the observed pixel belonging to the background or to the foreground. Section 4 describes a five-dimensional feature space that represents the certain pixel. In the final section, the experiment result is given to show the performance of proposed approach.

2 Kernel Density Estimation

In this part, the kernel density estimation modeling for background is introduced. The main objective for us is to capture the fast changing of local background scenes .We can take a certain kernel density function to determine the observation pixel probability of belonging to background scene. And this estimation is based on the chosen historic consecutive pixels' intensity.

Assume $x_1, x_2, x_3, \ldots\ldots x_n$, are a series of sample intensity over time for a certain pixel. Then, the probability density function that this pixel will have intensity value x_t at time t can be non-parametrically estimated using the kernel estimator K as follow [6]:

$$\Pr(\chi_t) = \frac{1}{n}\sum_{i=1}^{N} K(\chi_t - \chi_i) \tag{4}$$

And $x_1, x_2, x_3, \ldots\ldots x_n$ can also be a set of d-dimensional points in R^d, to represent the multiple characteristics as optical flow ,color of a pixel. So, the density estimation can be determined by (1),

$$\Pr(\chi_t) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - \chi_i) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{\|H\|^{1/2}} K(H^{-1/2}(x - \chi_i)) \tag{5}$$

H is a symmetric positive defined d*d matrix (bandwidth matrix), and K is a kernel function. The simplest approach is to use a fixed bandwidth matrix H for all the samples. A survey is provided in [6] of choosing bandwidth H. In most cases, two formulations of bandwidth estimator are considered: balloon estimator and sample-points estimator,

$$\Pr(x) = \frac{1}{n}\sum_{i=1}^{n} K_{H(x)}(x - x_i) \tag{6}$$

And sample-points estimator varies the bandwidth depending on the sample point, $H(x_i)$ is the bandwidth matrix at x_i.

$$\Pr(x) = \frac{1}{n}\sum_{i=1}^{n} K_{H(\chi_i)}(x - x_i) \tag{7}$$

Such a method is a compromise between complexity and the quality of approximation, but the use of variable bandwidth H can improve the accuracy of the estimated density. In some cases [6], in order to further reduce the computation, H is in the form of $H = h^2 I$ or $H = diag(h_1^2, h_2^2, \ldots h_d^2)$.

Kernel density estimation is also called Parzen Window in pattern recognition, for estimator K, square window function , triangular window function , Epanechnikov function are employed. Typically, Normal kernel function(Gaussian kernel) is widely used .

$$\Pr(x) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d}\frac{w_i}{\sqrt{2\pi\sigma_j^2}}e^{-\frac{1}{2}\frac{(x_{t_j}-x_{i_j})^2}{\sigma_j^2}} \tag{8}$$

3 Classify the Background and Foreground Pexel and Supress the Illumination Clutter

Using this probability function to estimate the pixel is considered as a foreground pixel if Pr(xt) < th where the threshold th is a global threshold over all the image that can be adjusted to achieve a desired percentage of false positives, as mentioned in [7].

$$M(i,j) = \begin{cases} 1, \Pr < th \\ 0, \Pr > th \end{cases} \tag{9}$$

Practically, the probability estimation equation above can be calculated in a very fast way using pre-calculated tables for the kernel function values given by the intensity value difference $(x_t - x_i)$.

After determining the background probability Pr by equation (8), we can decide the pixel whether it is belonging to background or foreground by (9).

In order to determine the threshold th, any available prior information about the characteristics of road, illumination of vehicle's lamp, and shadow should be utilized.

Another factor should be considered is to guarantee a false-alarm rate which is less than α_f requires that the threshold th should be set as:

$$\int_{\Pr(x)<th} \Pr(x)dx < \alpha_f \tag{10}$$

The sample points need to be updated when a new frame is coming, there are two alternative mechanisms to update the background[7]:

1) Selective Update: add the new sample to the model only if it is classified as a background sample.
2) Blind Update: just add the new sample to the model.

The first method enhances the detection for the target, because the target points are not added into the background sample points. The second method is simpler than the first one, but the problem with this method is that any incorrect detection decision will result in persistent incorrect detection result.

We can make a compromise based on the two methods for updating. A method of selective updating with background weight value is employed based on the normal updating sequences, which means that to give the new background point a larger weight value w_i when adding it into the estimation sample points.

Then the kernel density estimation with a background weight value w_i is as the equation (8).

Illumination suppression is on the basis of robust background model mentioned in the second and third section. The area of illumination is larger than normal noises, median filter and traditional mathematical morphological operator cannot eliminate this kind of area. So, prior information about it should be utilized to help with detecting the illumination area.

Fig. 2. The area figured out is the typical illumination area in the night environment

As is shown above the illumination area has some typical characteristics like it has a higher intensity and it is near the motion vehicle. So we can detect this area whose intensity ranges in certain area (β_1, β_2) and pixels in this area cannot be classified as background. Illumination area can be eliminated based on the following conditions:

$$M' = \{ \begin{matrix} 0; M(x,y) \cap (T_s \le \dfrac{I_l^s}{\underset{j=1...M}{med}(S^s)} \le T_s^2) \cap ((I_l^d - \underset{j=1...M}{med}(S_j^d)) \le T_d) \\ M_l; others \end{matrix} \tag{11}$$

Med(*) is the operator of getting the median value of the sample, and Ts is the parameter for the environment illumination intensity, Td is determined by experiments and historic observation data.

4 Feature Space of Pixel

We utilize five features to build the five-dimensional feature space, two for the space information, and three for normalized color representation. Traditional RGB model for color space has its demerits, especially when illumination slowly changes, and after normalization[8]:

$$r = R/S,\ g = G/S\ ,\ I = S/3 \quad \text{where } S = R+G+B$$

We use normalized r, g , I, to represent the color space , and under certain conditions , it is invariant to changing in illumination.

The covariance $\sum_i$ of an observation xi may be estimated from the covariance of the components, then the expression for the covariance matrix may be derived:

$$\sum_i = \begin{bmatrix} \Sigma_{r,g} & 0 & 0 \\ 0 & \sigma_i & 0 \\ 0 & 0 & \Lambda_f \end{bmatrix} \tag{12}$$

5 Experiment Result

The purpose of using KDE model is to build a robust background model for motion detection in the traffic scene in visual surveillance system. In the night circumstance of traffic scene, background modeling is difficult because illumination from vehicle lamp can easily bring clutters to background and the illumination may be detected as foreground area.

Under Kernel Density Estimation model, by taking adaptive threshold of calculated probability, background lamp illumination in small region can be suppressed. After processing of median filter and mathematical morphological operator, we can obtain a binary image to be adaptive to some cluster of background scene.

As is shown in the Fig.3 below, illumination change of lamp results in some detection errors that classify the light illumination into foreground scene, however, using KDE model together with the feature space that is mentioned above , detection error is reduced based on the background subtraction framework .

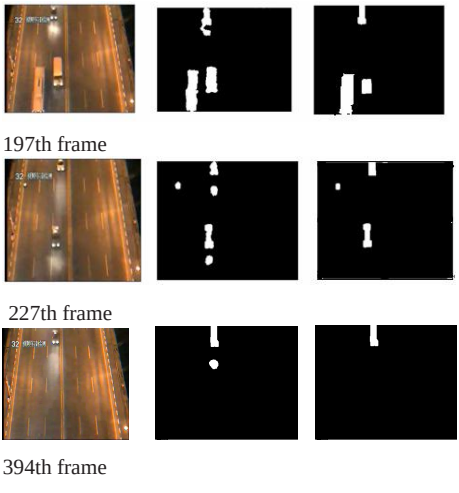


Fig. 3. The left row shows the original image sequences, and the middle row shows the detection result by using inter-frame difference method, the right row shows the detection result by using background subtraction based on KDE model for background (with the same Mathematical Morphological dilation and erosion)

In order to better validate the proposed technique, we also test our algorithm with another traffic sequences with waving trees, clutters to background in these images include waving trees and bushes, and environment illumination changes. The result in Fig.4 demonstrates the detection for moving car in the path. A typical traffic surveillance scenario as Fig.4 is challenging due to the vigorous motion of the tree (in the left) and the bushes (in top right corner), the aim for detecting the moving car in the top of he image sequences (from left to right) is disturbed by the waving clutters. As is shown in Fig.4, our algorithm shows the detection result in this situation. Again, our approach with KDE background modeling is able to deal with these sequences without clutter noises.

For the image sequences in Fig.5, the environment illumination changes continually and evidently when the walking person passes by. To test our algorithm with these sequences can assess the robustness of proposed algorithm for outdoor visual surveillance with environment illumination changes. Such illustration in Fig.5 with the result in Fig.3 shows the robustness character of our approach in outdoor visual surveillance for both day and night environments.

Fig. 4. Detection for a traffic sequence consisting waving trees and bushes. The upper images are the original ones, the other sequences are the detection results for moving car with our algorithm.

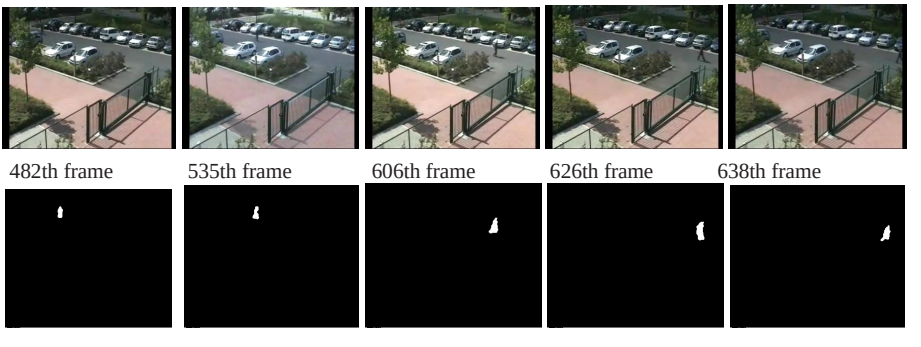


Fig. 5. Detection for a passenger consisting illumination change. The upper images are the original ones, the other sequences are the detection results for a walking person with our algorithm.

6 Conclusion

An approach for motion detection based on KDE model and multiple dimensional feature space is proposed in this paper, which can effectively suppress the background clutters like illumination from the lamp, environment illumination changes and other noises. Detection false is reduced and the system can satisfy the request for real-time processing. However, the demerit is that the performance is poor when the vehicles' occlusion appears. Such demerit is the key challenge for future work.

Acknowledgments. This work is supported by the National Basic Research Program of China(973 Program)(No.2007CB714107), and the National Natural Science Foundation of China(NSFC)(No.50679098).

References

1. Wren, C.R., Azarbayejani, A., Darrell, T.J., Pentland, A.P.: Pfinder: Real-time Tracking of the Human Body. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 780–785 (1997)
2. Friedman, N., Russell, S.: Image Segmentation in Video Sequences: A Probabilistic Approach. In: Thirteenth Conference on Uncertainty in Artificial Intelligence(UAI) (1997)
3. Grimson, W.E.L., Stauffer, C., Romano, R., Lee, L.: Using Adaptive Tracking to Classify and Monitor Activities in a Site. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, Santa Barbara, CA (1998)
4. Mittal, A., Paragios, N.: Motion-based Background Subtraction Using Adaptive Kernel Density Estimation. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition (2004)
5. Sheikh, Y.: Bayesian Modeling of Dynamic Scenes for Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(11) (2005)
6. Elgammal, A., Duraiswami, R., Davis, L.: Probabilistic Tracking in Joint Feature-Spatial Spaces. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition (2003)
7. Elgammal, A., Harwood, D.: Non-parametric Model for Background Subtraction. In: Proc. European Conference on Computer Vision (2000)
8. Elgammal, A., Harwood, D., Davis, L.: Background and Foreground Modeling Using Non-Parametric Kernel Density Estimation for Visual Surveillance. In: Proc. IEEE (2002)

Fingerprint Scaling*

Chun-Xiao Ren, Yi-Long Yin**, Jun Ma, and Hao Li

School of Computer Science and Technology, Shandong University, Jinan, 250100, China
{renchunxiao,ylyin,majun,li_hao}@sdu.edu.cn

Abstract. The problem of fingerprint scaling has received limited attention in the literature. Fingerprint scaling is an important issue for fingerprint sensor interoperability. However, no systematic study has been conducted to ascertain its effect on fingerprint systems. In this paper, a fingerprint scaling scheme using the average inter-ridge distance is presented. At first, the average inter-ridge distances of two fingerprint images acquired with two different fingerprint sensors are estimated respectively. Then the images are zoomed according to scale of the two average inter-ridge distances. Experiments have shown that the proposed fingerprint scaling method has good performance.

Keywords: Fingerprint, scaling, sensor interoperability, inter-ridge distance.

1 Introduction

Fingerprint recognition, as the most popular method in biometric authentication at present, has been used in identifying individuals for a long time [1]. Nowadays, most fingerprint systems aim at comparing images originated from the same sensor. In some cases, the automated fingerprint identification systems (AFISs) are trained and used on data obtained using the same sensor. However, the AFISs' ability to act on data from other sensors is restricted to a certain extent. This limitation baffles the use of multiple sensors with different characteristics in a single fingerprint system [2].

Fingerprint sensor interoperability means the ability of a fingerprint system to compensate for the variability introduced in the data of the same person due to employing different sensors [3]. The variations induced from the raw images owing to differences in resolution, scanning area, sensing technology, etc. impact the features extracted from the fingerprint images and propagate into the stored templates.

Fingerprint scaling refers to the procedure of fingerprint recognition to unify sizes of fingerprint obtained using the different sensor. Because the fingerprint's size in the images originated from the same sensor is similar, in fact, this problem does not exist in traditional AFISs. Rotation and translation are two facets considered in traditional AFISs.

* The work is supported by Shandong Province Science and Technology Development Program under Grant No. 2005GG3201089, Reward Fund for Outstanding Young and Middle Aged Scientists of Shandong Province under Grant No. 2006BS01008, and Shandong Province High Technology Independent Innovation Project under Grant No. 2007ZCB01030.

** Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 474–481, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Fingerprint scaling is an important issue for fingerprint sensor interoperability. Effective fingerprint scaling can not only deal with the variability introduced in the raw images due to differences in resolution, but also improve the reliability of matching considerably. Most fingerprint systems are limited in their ability to compare images originated from two different sensors resulting in poor inter-sensor performance [4]. Therefore, some scanners based on optical, capacitive, piezoelectric, thermal or ultrasonic principles are permitted to acquire fingerprint images [5]. The sensing area of these scanners may vary greatly from a few square millimeters to a few square inches. The resolution of the acquired image can vary anywhere between 250 dpi (e.g., Miaxis's FPR-361) and 512 dpi (e.g., Fang Yuan's U.are.U 4000); scanners that acquire 1000 dpi images (e.g., Aprilis' HoloSensor) are also available in the market.

It has been noted above that the fingerprint scaling is important for fingerprint identification. To the best of the authors' knowledge, however, the estimation of fingerprint scaling has received limited attention in the literature. Ross [3] studied the effect of matching fingerprints acquired with two different fingerprint sensors, resulting in a significant drop of performance. Alonso [5] studied the effect of matching two signatures acquired with two different Tablet PCs, resulting in a drop of performance when samples acquired with the sensor providing the worst signal quality are matched against samples acquired with the other sensor.

We present a fingerprint scaling method that constructs a relation between two images acquired using different sensors by utilizing computing the average inter-ridge distance. At first, the average inter-ridge distances of two fingerprint images are estimated respectively. Then the images are zoomed according to scale of two average inter-ridge distances.

This paper is organized as follows. Section 2 discusses fingerprint scaling method. Section 3 describes experiment procedures and presents some experimental results. Section 4 summarizes this paper and gives our some future work.

2 Fingerprint Scaling

The fingerprint scaling scheme involves two stages: estimating the average inter-ridge distances and zooming the fingerprint images, as illustrated by Fig.1. During estimation, the average inter-ridge distances are reckoned by a certain method. During zooming, the scale parameter is calculated and the fingerprints are unified to a similar size.

2.1 Estimating Average Inter-ridge Distance

The first step of the fingerprint scaling is estimation of the average inter-ridge distances in two fingerprint images originating from different sensors. Average inter-ridge distance is an important characteristic of fingerprint and is often used in fingerprint enhancement and ridge extraction procedures [1]. In this paper, the algorithm of average inter-ridge distance estimation transforms image using discrete Fourier transform (DFT) at first, and then estimates region through discrete information entropy, finally calculates average inter-ridge distance by weighted Euclidean distance.

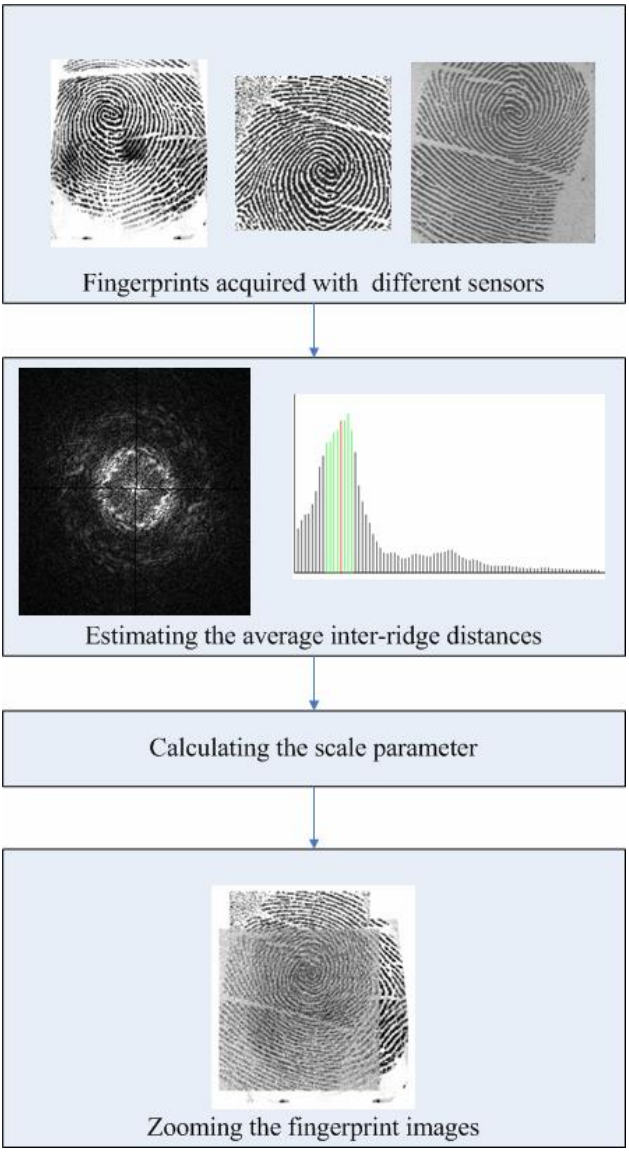


Fig. 1. The fingerprint scaling scheme

Fingerprint Image Analysis with Spectrum. Spectral analysis method is a typical method of signal processing in the frequency field. It transforms the representation of fingerprint images from the spatial field to the frequency field. It is a traditional method for average inter-ridge distance estimation in fingerprint images. In this method, we use discrete Fourier transform to analyze fingerprint images.

If $g(x,y)$ is the gray-scale value of the pixels with coordinates x, y $\in \{0,\ldots,N-1\}$ in an $N \times N$ image, the DFT of $g(x,y)$ is defined as fellow:

$$G_{(u,v)} = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} g_{(x,y)} e^{-2\pi j N\langle (x,y)(u,v)\rangle}$$

$$= \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} g_{(x,y)} \left(\begin{array}{l} \cos\left(-\frac{2\pi}{N}\langle (x,y)(u,v)\rangle\right) + \\ j\sin\left(-\frac{2\pi}{N}\langle (x,y)(u,v)\rangle\right) \end{array} \right) \quad (1)$$

Where j is an imaginary unit, $u, v \in \{0,\ldots,N-1\}$, and $\langle (x,y)(u,v)\rangle = xu + yv$ is the vector dot product; $G_{(u,v)}$ is obviously complex. Let $|G_{(u,v)}|$ denote the magnitude of $G_{(u,v)}$, theoretically speaking,

$$|G_{u,v}| = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} \sqrt{\begin{array}{l} \left(g_{(x,y)}\cos\left(-\frac{2\pi}{N}\langle (x,y)(u,v)\rangle\right)\right)^2 \\ +\left(g_{(x,y)}\sin\left(-\frac{2\pi}{N}\langle (x,y)(u,v)\rangle\right)\right)^2 \end{array}} \quad (2)$$

$|G(u,v)|$ is also called the coefficient. It represents the periodic characteristics of point u, v. The dominant period of signals in an area can be determined by analyzing the distribution of values of $|G(u,v)|$.

The whole procedure of average inter-ridge distance estimation with the spectral analysis method relies on a radial distribution function $Q(r)$ [7] defined as fellow:

$$Q(r) = \frac{1}{\#C_r}\sum_{(u,v)\in C_r} |G_{(u,v)}| \quad (3)$$

Where $0 \le r \le \sqrt{2}(N-1)$, Cr represents the set of coordinates u, v that satisfy $\sqrt{u^2+v^2} \approx r$, and $\#Cr$ is the number of elements of Cr. Based on [7], $Q(r)$ denotes distribution intensity of the signal whose period is N/r in an $N \times N$ image, and the value of r corresponding to the maximum of $Q(r)$ is the incident times of dominant signal in this area.

Estimating Region Using Entropy. Because the distribution is complicated, it is difficult to find reasonable frequency. Using entropy, we can locate the region which contains energy with largest density. In the DFT image, it equates getting the girdle with highest luminance.

Shannon information entropy theory is one of measurements to explore the distribution of energy [8]. Given a discrete probability distribution:

$$[X, P_x] = [x_k, p_k | k = 1,2,\ldots,K] \quad (4)$$

The information entropy of discrete random variable X is defined as:

$$H(X) = -\sum_{k=1}^{K} p_k \log p_k \tag{5}$$

For discrete random variables, larger entropy means larger uncertainty which depends on the range and uniformity of the distribution. Region with larger *H(x)* contains more information. So we can define density of information entropy for discrete random X as follow:

$$\overline{H(X)} = H(X)/k \tag{6}$$

Using $\overline{H(X)}$, we can get the region with largest density of entropy.

Frequency Computation Using Weighted Euclidean Distance. It is obviously that we have to find the most representative frequency in the girdle. Using weighted Euclidean distance, we can give a criterion to represent the capability of representative.

Weighted Euclidean distance [9] is defined as follow:

$$d_k(i,j) = \sqrt{\sum_{k=i}^{j} p_k \left(x_{ik} - x_{jk}\right)^2} \tag{7}$$

Where p_k is the energy, x_{ik} and x_{jk} are the distance from i to k and j to k respectively.

The ridge frequency with smallest $d_k(i,j)$ is the final result.

$$F = \left[\min d_k(i,j) \middle| k = i,\ldots,j\right] \tag{8}$$

2.2 Scaling Fingerprint Image

Suppose the set of fingerprints acquired from the same finger with different fingerprint sensors is represented as

$$E = \left\{F^i \mid i = 1,2,...,m\right\} \tag{9}$$

where *m* is the number of fingerprints and F^i is the *i*th fingerprint. We can calculate the average inter-ridge frequency $f(F^i)$ using the method described in section 2.1. Zooming the fingerprint F^i, it is obviously that there is linear amplification between F^i and F^{i*} which is the fingerprint image after zooming.

$$\frac{f\left(F^{i*}\right)}{f\left(F^i\right)} = \frac{s\left(F^{i*}\right)}{s\left(F^i\right)} \Rightarrow s\left(F^{i*}\right) = \frac{f\left(F^{i*}\right)}{f\left(F^i\right)} \times s\left(F^i\right) \tag{10}$$

where $s(F^i)$ is the size of fingerprint image.

Average inter-ridge distance of the same finger is stable correspondingly, so it is easy to see that we should scale the fingerprints of E to unify their average inter-ridge frequency. If the fingerprint image after zooming, F^{p*}, has the equal average inter-ridge frequency with F^q, viz.

$$f\left(F^{p*}\right) = f\left(F^{q}\right) \tag{11}$$

The size of F^{p*} will be

$$s\left(F^{p*}\right) = \frac{f\left(F^{q}\right)}{f\left(F^{p}\right)} \times s\left(F^{p}\right) \tag{12}$$

where $p, q \in \{1,2,...,m\}$.

3 Experiment

In order to evaluate the method, we have done experiment on man-made database acquired with two different fingerprint sensors.

We made 2 groups of images for experiment. There are 50 images in every group, and 100 images totally are used in experiment. Images in experimental sets were transformed from FVC 2000 DB1 and DB2. The transformation contains scaling, rotation, and translation by random within a certain range. Fig.2 (a), (b) show original and transformed images with their parameter as scaling(73%), rotation(-19°) and translation(-16%,-3%) respectively.

(a) Original image:DB2_5 (b) Image after transformation

Fig. 2. Man-made database

The experiment proceeds as follows:

1. The average inter-ridges of original image and transformed image, F^i and F^{i*} respectively, are calculated using the method described in section 2.1.
2. The predictive scaling value *p(i)* is computed.

$$p(i) = \frac{f(F^{i*})}{f(F^{i})} \tag{13}$$

3. The departure of predictive scaling value *d(i)* is computed.

$$d(i) = \left| \frac{p(i)}{r(i)} - 1 \right| \tag{14}$$

where *r(i)* is the value of real parameter of scaling.

The experimental result, departure of predictive scaling value, for the Database 1 and 2 of FVC2000, using our method is 11.98%.

4 Summary and Future Work

The need for fingerprint scaling is pronounced because of the widespread deployment of fingerprint systems in various applications and the request of sensor interoperability. In this paper, we have illustrated the impact of estimating fingerprint scaling value on the sensor interoperability of a fingerprint system. Fingerprint scaling is an important issue for fingerprint sensor interoperability. However, no systematic study has been conducted to ascertain its effect on fingerprint systems.

In this paper, a fingerprint scaling algorithm using the average inter-ridge distance is presented. At first, the average inter-ridge distances of two fingerprint images are estimated respectively. Then the images are zoomed according to scale of two average inter-ridge distances. Experiments have shown that the proposed fingerprint scaling method has good performance.

Currently, the man-made database is used to evaluate the method. We are working on real-world database that acquired with 3 different fingerprint sensors to enhance the sensor interoperability of a fingerprint system.

References

1. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, New York (2003)
2. Ross, A., Nandakumar, K., Jain, A.K.: Handbook of Multibiometrics. Springer, Heidelberg (2006)
3. Ross, A., Jain, A.: Biometric Sensor Interoperability: A Case Study in Fingerprints. In: Maltoni, D., Jain, A.K. (eds.) BioAW 2004. LNCS, vol. 3087, pp. 134–145. Springer, Heidelberg (2004)
4. Ross, A., Nadgir, R.: A Calibration Model for Fingerprint Sensor Interoperability. In: Proc. of SPIE Conference on Biometric Technology for Human Identification III, Orlando, USA, p. 62020B-1 – 62020B-12 (2006)
5. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2002: Fingerprint Verification Competition. In: Proceedings of the International Conference on Pattern Recognition (ICPR), Quebec City, Canada, pp. 744–747 (2002)

6. Alonso-Fernandez, F., Fierrez-Aguilar, J., Ortega-Garcia, J.: Sensor Interoperability and Fusion in Signature Verification: A Case Study Using Tablet PC. In: Li, S.Z., Sun, Z., Tan, T., Pankanti, S., Chollet, G., Zhang, D. (eds.) IWBRS 2005. LNCS, vol. 3781, pp. 180–187. Springer, Heidelberg (2005)
7. Kovacs-Vajna, Z.M., Rovatti, R., Frazzoni, M.: Fingerprint Average Inter-ridge Distance Computation Methodologies. Pattern Recognition 33(1), 69–80 (2000)
8. Zuo, J., Zhao, C., Pan, Q., Lian, W.: A Novel Binary Image Filtering Algorithm Based on Information Entropy. In: The Sixth World Congress on Intelligent Control and Automation (WCICA), vol. 2, pp. 10375–10379 (2006)
9. Gower, J., Legendre, P.: Metric and Euclidean Properties of Dissimilarities Coefficients. J. Classification 3, 5–48 (1986)

A Novel Classifier Based on Enhanced Lipschitz Embedding for Speech Emotion Recognition

Mingyu You[1], Guo-Zheng Li[1,2,⋆], Luonan Chen[2], and Jianhua Tao[3]

[1] School of Computer Engineering and Science, Shanghai University, Shanghai 200072, China
[2] Institute of System Biology, Shanghai University, Shanghai 200444, China
[3] National Laboratory of Pattern Recognition, Chinese Academy of Sciences Beijing 100080, China
gzli@shu.edu.cn

Abstract. The paper proposes a novel classifier named ELEC (Enhanced Lipschitz Embedding based Classifier) in the speech emotion recognition system. ELEC adopts geodesic distance to preserve the intrinsic geometry of speech corpus and embeds the high dimensional feature vector into a lower space. Through analyzing the class labels of the neighbor training vectors in the compressed space, ELEC classifies the test data into six archetypal emotional states, i.e. neutral, anger, fear, happiness, sadness and surprise. Experimental results on a benchmark data set demonstrate that compared with the traditional methods, the proposed classifier of ELEC achieves 17% improvement in average for speaker-independent emotion recognition and 13% for speaker-dependent.

Keywords: Lipschitz Embedding, Speech Emotion Recognition.

1 Introduction

Emotion recognition is a challenging problem in the intelligent computation field. Accurate emotion recognition from speech signals will be applied to areas of human-machine interaction, entertainment, tele-learning, preventive medicine, consumer relations, etc. [1]. A general process of speech emotion recognition is formulated as: 1) extracting acoustic features 2) reducing the feature dimensionality and 3) recognizing emotions with a classifier.

Dimensionality reduction is vital to speech emotion recognition, which contains two categories, i.e. feature extraction and feature selection. Feature selection methods like Sequential Forward Search [2] and Sequential Floating Forward Search [3] need heavy computation to evaluate all the feature subsets. And only part of the classification information remains. Therefore, feature extraction methods like PCA(Principal Component Analysis), LDA(Linear Discriminant Analysis) and MDS(Multidimensional Scaling) are frequently mentioned. Chuang and Wu [4] adopted PCA to select 14 principle components in the analysis of emotional

⋆ Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 482–490, 2008.
© Springer-Verlag Berlin Heidelberg 2008

speech. LDA and MDS have also been employed to reduce the feature dimensionality for emotion recognition [5]. Though widely used for their simplicity, PCA, LDA and MDS are limited by their underlying assumption that data lies in a linear subspace. In fact, speech points are shown to reside on a nonlinear submanifold [6]. So, some nonlinear techniques have been proposed to discover the structure of manifolds, e.g. Isomap [7] and LLE(Locally Linear Embedding) [8]. Lipschitz embedding [9] is another nonlinear dimensionality reduction method which works well when there are multiple clusters in the input data [10]. But they yield maps that are defined only on training vectors, and how to embed the novel test data into the map remains unclear. To handle the problem, You et al. [11] proposed ELE (enhanced Lipschitz Embedding), which exhibited better performance than the previous methods.

Compressed feature vectors are classified or recognized by different classifiers. HMM(Hidden Morkov Model) prefers temporal features embodying the instantaneous information. SVM(Support Vector Machine), NN(Neural Network), Bayesian Network are all discriminative classifiers suitable for statistic features. SVM is widely studied by researchers [12],where linear and nonlinear kernel functions enable it excellent performance.

In this paper, a classifier named ELEC is developed to analyze the intrinsic manifold of emotional speech, and recognize the emotional states. Furthermore, it is compared with different combinations of methods PCA, LDA, SFS, Isomap, LLE and the SVM classifier in terms of prediction accuracy.

The rest of the paper is organized as follows. Section 2 gives a brief description of the system and the details of ELEC. Experimental results are provided and discussed in Section 3. Section 4 concludes the paper and future work.

2 A Novel Classifier Based on Enhanced Lipschitz Embedding

2.1 A Speech Emotion Recognition Framework

The framework of a speech emotion recognition system with our proposed classifier ELEC is illustrated in figure 1. Firstly, high dimensional acoustic feature vectors are extracted from the speech utterances. Then, speech data is split into training and test parts according to the 10-fold cross validation. After ELEC is trained by the training vectors, it is used to recognize the emotional state of the test data.

2.2 A Training Algorithm of ELEC

ELEC is based on enhanced Lipschitz embedding, which is introduced firstly. A Lipschitz embedding is defined in terms of a set R ($R = \{A_1, A_2, \ldots, A_k\}$), where $A_i \subset S$ and $\bigcup_{i=1}^{k} A_i = S$. The subset A_i is termed as the reference set of the embedding. Let $d(o, A)$ be an extension of the distance function d from object o to a subset $A \subset S$, such that $d(o, A) = \min_{x \in A} d(o, x)$. An embedding with respect to R is defined as a mapping F such that $F(o) = (d(o, A_1), d(o, A_2), \ldots, d(o, A_k))$.

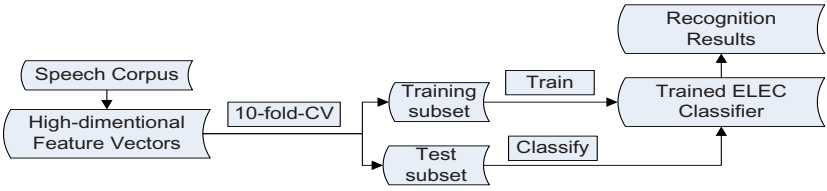


Fig. 1. A Framework of Speech Emotion Recognition with ELEC

In other words, Lipschitz embedding defines a coordinate space where each axis corresponds to a subset $A_i \subset S$ and the coordinate values of object o are the distances from o to the closest element in each A_i. In our algorithm, the speech corpus is divided into six subsets $\{A_1, A_2, \ldots, A_6\}$ according to six emotional states.

The distance function d in Lipschitz embedding reflects the essential structure of data set. Due to the nonlinear geometry of speech manifold, the distance between the speech examples in space is measured by the geodesic distance.

In ELEC, a graph G is set up, whose edges only connect the points representing neighbor examples in the original space. Matrix M is used to indicate the information in the graph G. The element m_{ij} of matrix M is initialized as:

$$m_{ij} = \sqrt{\sum_{\alpha=1}^{Hd} (i_\alpha - j_\alpha)^2} \tag{1}$$

where element m_{ij} stands for the Euclidean distance from point i to j. $i = [i_1, i_2, \cdots, i_{Hd}]$ and $j = [j_1, j_2, \cdots, j_{Hd}]$ are data points in the Hd-dimensional feature space. Hd is the dimensionality of the original space. It is set to 64 in this paper, as described in section 3.1.

A training algorithm of the ELEC classifier is illustrated in Algorithm 1, whose basic idea is that the coordinate values of a object o are the geodesic distances from o to the closest elements in each emotional state. Function Findmin in step 7 of Algorithm 1 finds the minimal sum of matrix M's elements from index i to j. o's coordinate values can be retrieved from matrix M like:

$$o^\gamma = \min_{\mu \in A_\gamma} m_{o\mu} \tag{2}$$

where o^γ is the o's γth coordinate value. $m_{o\mu}$ is an element of matrix M, standing for the distance between o and μ. μ is an object from the emotional state set A_γ.

2.3 A Test Algorithm of ELEC

In the test phase, how to embed the new coming data into the lower space is a tough task. The matrix M is constructed on the training data, and it's impossible to reconstruct it combining new coming data. Based on the constructed M, a

Algorithm 1. A Training Algorithm of ELEC

1: INPUT: a training data set $\boldsymbol{Tr} \subset \boldsymbol{\mathcal{R}}_N^{Hd}$, with its emotion label vector $\boldsymbol{La} \subset \boldsymbol{\mathcal{R}}_N$, where $\boldsymbol{N}$ is the number of training data and $\boldsymbol{Hd} = 64$ is the dimension of acoustic feature vector in original space.
2: Set up matrix $\boldsymbol{M}$ with dimension $\boldsymbol{N} \times \boldsymbol{N}$, using Equation (1). $\boldsymbol{N}$ is 3240 in this paper.
3: **if** $\boldsymbol{m}_{ij}$ is not among the $\boldsymbol{k}$ smallest elements of $\boldsymbol{m}_{io}$, $\forall_{o \in N}$ **then**
4: Set $\boldsymbol{m}_{ij}$ to $\boldsymbol{INF}$, $\boldsymbol{INF}$ is an infinite number.
5: **end if**
6: **if** $\boldsymbol{m}_{ij} == \boldsymbol{INF}$ **then**
7: $\boldsymbol{m}_{ij} = \mathbf{Findmin}(\boldsymbol{m}_{io_1} + \ldots + \boldsymbol{m}_{o_{l-1}o_l} + \ldots + \boldsymbol{m}_{o_pj})$
8: **end if**
9: Get the coordinate values $\boldsymbol{ELETr}_{\boldsymbol{i}}^{\gamma}$ for each training data $\boldsymbol{i}$, using Equation (2).
10: OUTPUT: a embedded set of training data $\boldsymbol{ELETr} \subset \boldsymbol{\mathcal{R}}_N^{\boldsymbol{Ld}}$, $\boldsymbol{Ld} = 6$ is the dimension of the embedded low dimensional space.

new approach is embedded into ELEC to compute the coordinate values of test data te in the lower space. te's geodesic distance to each subset is the minimal sum of "short hops" to neighbor point and the geodesic distance of the neighbor. In other words, te makes the shortest pathes to subsets through its neighbors. We can prove the idea in theorem 1. te's coordinate values can be calculated by:

$$te^{\gamma} = \min \sum_{\alpha=1}^{k} (d_{te\alpha} + ELETr_{\alpha}^{\gamma}) \tag{3}$$

where te^{γ} represents the γth coordinate value of te, $d_{te\alpha}$ stands for the distance between te and α. $ELETr_{\alpha}^{\gamma}$ is the training data α's γth coordinate value in the compressed space. ELEC classifies the test data into the class that most of its neighbors in the compressed space belong to.

Theorem 1. *Test data's geodesic distance to each subset is the minimal sum of the distance from test data itself to a neighbor point and the geodesic distance of the neighbor.*

Proof. Step 1: Test data's class label is unknown, so the geodesic path nearest to each known set must pass through some training points in the space.
Step 2: As geodesic paths always go through the neighbor points, test object te's geodesic path will start from one of the edges connecting te and its neighbors.
Step 3: For example, te's geodesic path to set i firstly goes through the neighbor point o. Then the following trace must be same with the geodesic path of o to set i.
Step 4: In Step 3, if there is another path shorter, the geodesic path of point o to set i will also be changed. So, the conclusion in Step 3. can be obviously proved by contradiction. te's geodesic distance to each set is the sum of the distance to its neighbor point and the geodesic distance of the neighbor. □

Algorithm 2. A Test algorithm of ELEC

1: INPUT: a embedded set of training data $\boldsymbol{ELETr} \subset \boldsymbol{\mathcal{R}}_N^{Ld}$, with it emotion label vector $\boldsymbol{La} \subset \boldsymbol{\mathcal{R}}_N$. a test data set $\boldsymbol{Te} \subset \boldsymbol{\mathcal{R}}_N^{Hd}$.
2: Find test data $\boldsymbol{Te_i}$'s $\boldsymbol{k}$ nearest neighbors $\{\boldsymbol{n_1}, \boldsymbol{n_2}, \ldots, \boldsymbol{n_k}\}$ in the training data set. The distances between $\boldsymbol{Te_i}$ and the neighbors are $\{\boldsymbol{d_{te1}}, \boldsymbol{d_{te2}}, \cdots, \boldsymbol{d_{tek}}\}$, respectively.
3: Get the coordinate values $\{\boldsymbol{ELETr_\alpha^1}, \boldsymbol{ELETr_\alpha^2}, \ldots, \boldsymbol{ELETr_\alpha^{Ld}}\}$ of the $\boldsymbol{\alpha}$th neighbor from matrix $\boldsymbol{M}$.
4: Compute the test data' coordinate values in the compress space, using Equation (3).
5: Calculate the class label frequency of the test data $\boldsymbol{Te_i}$'s $\boldsymbol{k}$ nearest neighbors, based on Equation (4).
6: Classify the test data $\boldsymbol{Te_i}$ as: $\boldsymbol{TLa_{Te_i}} = \arg\max \boldsymbol{N}_i$.
$\boldsymbol{Te_i}$ is classified into the class that most of its $\boldsymbol{k}$ nearest neighbors belong to. If there is equal case, $\boldsymbol{Te_i}$ is given randomly into the maximal classes.
7: OUTPUT: recognized class label vector $\boldsymbol{TLa} \subset \boldsymbol{\mathcal{R}_N}$ of the test sample $\boldsymbol{Te}$

The test algorithm of the ELEC classifier is explained in Algorithm 2. Ld means the number of class labels which is six in this paper. The number of training data, which are test example te's neighbors, belonging to each class is calculated by:

$$N_i = \sum_{tr} (La_{tr} == i) \qquad \forall tr, te \in KNN \tag{4}$$

were La_{tr} represents the class label of the training data tr. N_i stands for the number of the training data whose class label is i. $tr, te \in KNN$ means that tr is among the k nearest neighbors of test data te.

In the next section, ELEC is set up in a speech emotion recognition system, and compared with other popular methods PCA, LDA, SFS, Isomap, LLE together with SVM.

3 Experiments

3.1 Speech Corpus

The speech database used in this paper are collected in National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences. It's an emotional speech corpus in Mandarin. The corpus is collected from four Chinese native speakers including two men and two women. Everyone expresses 300 sentences in six emotions. The total amount of sentences is 7200. The speech corpus is sampled at 16kHZ frequency and 16 bits resolution with monophonic Windows PCM format.

In this work, 48 prosodic and 16 formant frequency features are extracted, include: max, min, mean, median of Pitch (Energy); mean, median of Pitch (Energy) rising/ falling slopes; max, mean, median duration of Pitch (Energy) rising/ falling slopes; mean, median of Pitch (Energy) plateaux at maxima/

minima; max, mean, median duration of Pitch (Energy) plateaux at maxima/minima. If the first derivative of Pitch (Energy) is close to zero and the second derivative is positive, the point belongs to a plateau at a local minimum. If the second derivative is negative, it belongs to a plateau at a local maximum. Formant frequency features which are widely used in speech processing applications were also investigated. Statistical properties including max, min, mean, median of the first, second, third, and fourth formant were extracted.

In the experiment, speaker-independent and speaker-dependent emotion recognitions are both investigated within the same gender. 10-fold cross-validation method is adopted considering the confidence of recognition results.

3.2 Experimental Setting

In order to evaluate the classification results of ELEC, linear methods PCA and LDA, nonlinear techniques Isomap and LLE, feature selection by SFS, together with SVM are implemented for comparison. SVM has originally been proposed for binary classification. In the experiments, 15 one-to-one SVMs are combined into an MSVM (Multi-SVM), in which each SVM is used to distinguish one emotion from another. Final classification result is determined by all the SVMs with the majority rule. After the heavy tests of polynomial, radial basis function and linear kernels with different parameters, linear SVM (C=0.1) is chosen for its acceptable performance and simplicity.

In the implementation of Isomap and LLE, we reconstruct the distance matrix M when facing the novel test data. Although it costs a lot of computation time, it will help attain the best performance of Isomap and LLE. Comparison with those results gives a solid evaluation of ELEC.

In ELEC, $k = 10$ nearest neighbors are searched in constructing the matrix M. The impact of different k on the system performance is also investigated, through a series of experiments with $k = 5 : 5 : 100$, we find $k = 10$ makes an acceptable performance with relatively low computational cost.

3.3 Experimental Results

Table 1 demonstrates the comparative performance in speech emotion recognition. From Table 1, we can observe that:

1. In speaker-independent and speaker-dependent processes, ELEC comes up with the best performance in almost all of the emotional situations, especially angry and fear. The average improvement of the ELEC classifier is 17% in the speaker-independent system and 13% when speaker-dependent.
2. Nonlinear methods are not in a prominent position compared with other techniques. Within nonlinear methods, Isomap+SVM outperforms LLE+SVM. Besides, LLE+SVM conducts unbalanced performance when dealing with different emotions. In speaker-independent, LLE+SVM only attains 10.1% accuracy with the emotion fear, while it achieves about 67% with sad.
3. The Linear method LDA+SVM achieves 10% higher average accuracy than PCA+SVM.

Table 1. Comparative Results of ELEC and Different Methods in Speech Emotion Recognition

Emotion	PCA+SVM	LDA+SVM	Isomap+SVM	LLE+SVM	SFS+SVM	ELEC
Male/Speaker-independent						
Angry	43.5%	65.2%	48.4%	33.7%	51.0%	79.8%
Fear	33.8%	37.8%	35.8%	29.3%	50.2%	62.3%
Happy	32.2%	36.2%	32.7%	32.9%	34.5%	59.7%
Neural	59.3%	72.0%	73.3%	81.0%	64.8%	77.0%
Sad	47.2%	56.8%	43.7%	32.7%	49.5%	63.0%
Surprise	47.2%	53.7%	60.9%	59.9%	49.7%	75.2%
Average	43.9%	53.6%	49.1%	44.9%	49.9%	69.5%
Female/Speaker-independent						
Angry	51.5%	60.5%	44.3%	28.4%	44.7%	79.2%
Fear	33.0%	48.7%	37.8%	10.1%	43.2%	60.1%
Happy	26.7%	46.3%	38.7%	51.8%	37.5%	61.8%
Neural	45.2%	67.0%	54.4%	42.6%	63.5%	70.0%
Sad	62.3%	68.2%	63.9%	66.8%	62.5%	72.2%
Surprise	37.0%	40.8%	48.3%	27.7%	48.0%	53.6%
Average	42.6%	55.3%	47.9%	37.9%	49.9%	66.2%
Male/Speaker-dependent						
Angry	66.0%	85.0%	88.7%	58.5%	77.0%	91.7%
Fear	60.7%	59.3%	68.0%	58.0%	74.0%	75.0%
Happy	48.3%	68.7%	58.1%	45.2%	61.0%	74.0%
Neural	81.3%	84.3%	89.8%	89.2%	50.7%	85.0%
Sad	64.3%	53.0%	68.2%	60.0%	56.0%	66.0%
Surprise	47.7%	77.0%	63.8%	68.5%	80.0%	78.2%
Average	61.4%	71.2%	72.8%	63.2%	66.4%	78.3%
Female/Speaker-dependent						
Angry	63.0%	79.0%	58.5%	48.2%	72.0%	89.0%
Fear	40.3%	56.0%	42.7%	13.6%	67.0%	68.0%
Happy	44.7%	68.7%	22.5%	13.6%	48.0%	72.0%
Neural	71.7%	82.0%	43.7%	13.8%	51.3%	80.0%
Sad	59.3%	65.7%	68.3%	58.5%	69.3%	69.0%
Surprise	53.3%	67.3%	36.9%	68.5%	67.0%	70.0%
Average	55.4%	69.8%	45.4%	36.0%	62.4%	74.7%

4. Average classification accuracy of the speaker-dependent emotion recognition is about 10% higher than that of the speaker-independent. Classification accuracy of the male speaker is a little higher than that of female. Among the emotional states, classification accuracy of happiness is lower than other emotions in the speaker-independent system, while accuracy of happiness is comparable with the others in the speaker-dependent.

3.4 Discussions

Observations in the above section show ELEC is better than the other recognition methods, which is reasonable. ELEC preserves the intrinsic geometry of emotional speech corpus, and maintains the distribution properties of emotion sets. If we project the embedded results of ELEC to a visual space, points of the same emotional state are highly clustered around one plane, which highly encourage the emotion classification.

There are acoustic variations between different people, emotion recognition accuracy of speaker-dependent must be higher than that of speaker-independent. Women's facial expressions or body gestures convey more emotional information, multi-modal emotion recognition is necessary in the future. Unbalanced recognition rate will greatly reduce the systems robustness. Isomap and LLE, with

SVM, behave differently between the male and the female in speaker-dependent. On the other hand, ELEC seems stable. Nonlinear methods performance may not be better than that of the linear methods, although they require more complicated computation. It is shown that the method of preserving the geometry of data set is crucial in nonlinear approaches.

4 Conclusion and Future Work

A novel classifier ELEC is proposed and compared with different combinations of state-of-arts linear and nonlinear dimensionality reduction methods with SVM. ELEC comes up with the best performance when dealing with almost all of the basic emotions in both speaker-independent and speaker-dependent processes. Although LDA+SVM and Isomap+SVM achieve plausible results, both of them lack the balance of classification in each emotion, while ELEC does. Besides, the time consumption of Isomap+SVM is much higher than that of ELEC.

Although ELEC achieves satisfactory results, there are still some future works like parallelization to improve its running efficiency and feature selection embedding to improve its generalization performance.

Acknowledgements. This work was supported by NSFC (20503015) and STCSM (07DZ19726).

References

1. Picard, R.: Affective Computing. The MIT Press, Cambridge (1997)
2. Grimm, M., Kroschel, K., Narayanan, S.: Support Vector Regression for Automatic Recognition of Spontaneous Emotions in Speech. In: Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, pp. 1085–1088 (April 2007)
3. Lugger, M., Yang, B.: The Rrelevance of Voice Quality Features in Speaker Independent emotion recognition. In: Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 4, pp. 17–20 (April 2007)
4. Chuang, Z.: Motion Recognition Using Acoustic Ffeatures and Textual Content. In: Proceedings of IEEE International Conference on Multimedia and Expo., pp. 53–56 (2004)
5. Go, H., Kwak, K., Lee, D., Chun, M.: Emotion Recognition from the Facial Image and Speech Signal. In: Proceedings of SICE 2003 Annual Conference, vol. 3, pp. 2890–2895 (2003)
6. Jain, V., Saul, L.K.: Exploratory Analysis and Visualization of Speech and Music by Locally Linear Embedding. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, pp. 984–987 (2004)
7. Tenenbaum, J., Silva, V.d., Langford, J.C.: A Global Geometric Framework for Nonlinear Dimensionality Rreduction. Science 290, 2319–2323 (2000)
8. Roweis, S., Saul, L.: Nonlinear Dimensionality Reduction by Locally Linear Embedding. Science 290, 2323–2326 (2000)
9. Bourgain, J.: On Lipschitz Embedding of Finite Metric Spaces in Hilbert Space. Israel J. Math. 52, 46–52 (1985)

10. Chang, Y., Hu, C., Turk, M.: Probabilistic Expression Analysis on Manifolds. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 520–527 (2004)
11. You, M., Chen, C., Bu, J., Liu, J., Tao, J.: Emotional Speech Analysis on Nonlinear Manifold. In: Proceedings of IEEE International Conference on Pattern Recognition, vol. 3, pp. 91–94 (2006)
12. Hu, H., Xu, M., Wu, W.: Gmm Supervector Based Svm with Spectral Features for Speech Emotion recognition. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 413–416 (April 2007)

GPU Implementation of a Clustering Based Image Registration*

Seung-Hun Yoo, Yun-Seok Lee, Sung-Up Jo, and Chang-Sung Jeong**

Department of Electronics and Computer Engineering, Korea University,
Anam-Dong, Seongbuk-Gu, Seoul, Korea
{capstoney,ruhmreich,easyup,csjeong}@korea.ac.kr

Abstract. This paper presents GPU(Graphics Processing Unit) implementation of a clustering based image registration method. Since the image registration is an important process in image analysis tasks such as image restoration and image fusion, fast image registration can improve the overall application execution speed. Recently, the commodity GPU is being used in not only 3D graphics rendering but also in general-purpose computation due to an increase in the good price/performance ratio and hardware programmability as well as the huge computing power and speed of the GPU. We implemented clustering-based image registration method on GPU using only transformation of texture coordinations in vertex program and re-sampling in fragment program. Finally, GPU-based image registration speed up nearly 50 percent compared with CPU.

Keywords: Image registration, Clustering, GPU.

1 Introduction

Image registration is the process of determining the point-by-point correspondence between two images of a scene [1][2]. Image registration is an important process in image analysis tasks such as object recognition, multichannel image restoration and image fusion, fast image registration can be enhancement of overall application performance.

Image registration methods can be classified into two categories[2][3][4]:area-based and feature-based . Area-based methods directly matching image intensities without any structural analysis. On the other hand, feature-based methods use common features such as edges, corners and closed-boundary regions. Area-based methods usually adopt a window of points to determine a matched location using the correlation technique and feature-based methods find the feature correspondence using spatial relations of features or various descriptors of features.

* This research was supported by the Seoul R&BD Program, a program of MIC (Ministry of Information and Communication) of Korea, the ITRC(Information Technology Research Center) support program of IITA (Institute of Information Technology Assessment), a Brain Korea 21 project, Seoul Forum for Industry-University-Research Cooperation and a grant of Korea University.

** Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 491–497, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Clustering techniques presented by Stockman et al in [5], use spatial relations of features and estimate the transformation parameters through a voting process[2][6]. After feature detection, feature matching is carried out between all possible pairs of CPs(control points) from the reference and target image [1]. The parameters of the transformation which maps the points on each other are computed and represented as a point in the space of transform parameters. Correct matches tend to make a cluster while mismatches fill the parameter space randomly. The parameter values corresponding to the center of densest cluster are used mapping function parameters.

Graphic hardware was specially made for the graphic operation such as the mathematical computation, acceleration of image processing and signal conversion. Graphic Processing Units(GPUs), which improve 3D graphic accelerator, are the core that performs graphic computation handling [7][8]. This device very quickly accomplishes the works of the graphics. Recently, GPU's ability of the computation has been also rapidly advanced. Particulary, the data parallelism in typical image processing is very well suited for data stream based GPU's architecture [9][10].

In this paper, we present a GPU(Graphics Processing Unit) implementation of clustering based image registration method. We considered the geometric relationship between the images as a similarity transformation which allows translation, rotation and scale factors between views. For more easy implementation, we estimates the number of feature consensus pairs in the feature space according to change of parameter values conversion over the parameter space.

The organization of the paper is as follow: Section 2 gives an architecture and programming model of GPU; Section 3 discusses the registration process and describes how implement it using graphics hardware; Section 4 shows the experimental results of registering images taken from same scenes and Section 5 summarized our work.

2 Graphics Processing Unit

Commodity graphics hardware, known as Graphics Processing Unit or GPU, has seen incredible growth in terms of performance, programmability, and arithmetic precision. In comparison to today's CPUs, the GPU provides a reduced set of instructions that support simple control flow structures and several mathematical operations, many of which are devoted to graphics specific functions. In this section, we introduce an GPU architecture and programmability of GPU.

2.1 GPU Programming Model

The stream programming model is the basis for programming GPUs today [9]. In the stream programming model, all data is represented as a stream which define as an ordered set of data of the same data type and kernels are small program operating on streams. Kernels take a stream as input and produce a stream as output.

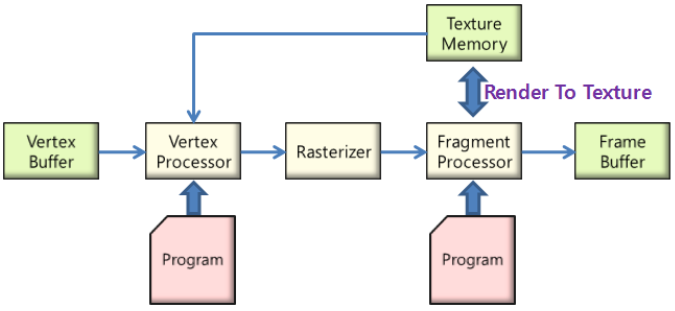


Fig. 1. Graphics Pipeline

All of today's commodity GPUs have followed common structure known as the graphics pipeline [7][9]. The graphics pipeline is traditionally structured as stages of computation connected by data flow between the stages. This graphics pipeline is similar to the stream and kernel abstractions of the stream programming model. A simplified current graphics pipeline is shown in Fig. 1. In hardware, each stage is implemented as a separate piece of hardware on the GPU.

This structure allows for three-level parallelism in high computational rates and throughput [9][10]. Basically, each stage may be able to run several tasks on different data at the same time (i.e. instruction-level and task-level parallelism). In addition, the pipeline is composed of multiple vertex and pixel/fragment processors, allowing for many data-parallel operations over the supplied set of input vertices and resulting pixels (i.e. data-level parallelism). For example, common hardware designs use 6-8 vertex processors and 32-48 pixel processors.

2.2 Programmable Hardware

As shown in Fig. 1, programmers can supply code for both the vertex and fragment processors [7][9][10][11]. Each program is typically limited to a fixed number of available input and output parameters, registers, constant values, and overall number of instructions. The input parameters to programs can be scalar values, two-, three-, or four-component vectors, or arrays of values (scalar or vector) that are stored in the texture memory of the graphics card. The GPU programming has the advantage and disadvantage in memory access sides. For example, a frame buffer memory can be display or can be written to texture memory, this work called RTT(Render To Texture) [7][12][13]. It is implemented direct feedback of GPU output to input without going back to CPU, so more efficient results is made. On the other hand, fragment processor can't use indirect memory addressing in writing operation and same memory as input and output.

Shading languages are directly targeted at the programmable stages of the graphics pipeline. The most common shading languages are Cg("C for graphics") [14], the OpenGL Shading Language, and Microsoft's High-Level Shading Language. Cg is a high-level shading language developed by NVIDIA, its syntax

and semantics are very similar to the C programming language. Our registration processes are implemented on GPU using Cg and OpenGL with graphics user interface.

3 The Registration Processes

3.1 Matching Using Clustering

The assumed geometrical model of clustering technique is the similarity transform. Similarity transform is the simplest model-it consists of translation, scaling and rotation[1][2].

$$u = s(xcos(\theta) - ysin(\theta)) + t_x \tag{1}$$

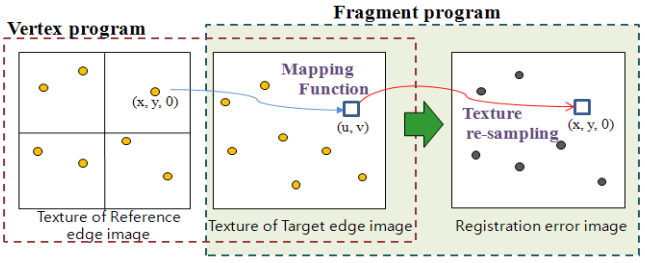
$$v = s(xsin(\theta) + ycos(\theta)) + t_y \tag{2}$$

In clustering, the transformation parameters are estimated through a voting process. For the similarity transform model, four 1-D accumulator arrays are used to estimate parameters. By using the possible ranges of the translational(t_x,t_y), rotational θ, and scaling s, coefficients of four parameters can be estimated. As a spatial difference between two(or more) aerial images is not large, the constraint of parameter range decreases computational complexity. In our implementation, we restricted the scaling difference between 0.5 and 2, rotational difference from -$\pi/6$ to $\pi/6$ and translational difference within 10% of the image size respectively.

Before execute clustering based feature matching, we extract feature points using canny edge detector having good detection and good localization in various situation [15]. Then, we count the number of matching pairs after varying the transformation parameters with respect to a predetermined quantity. This step is repeated several times and the parameters giving the maximum number of matching pairs are selected as the answer to the transformation. If the distance between matching pairs is less than a given value, the matching pairs are regarded as correct. Parameters with the maximum value of matched feature pairs set the mapping function parameters and image re-sampling and transformation are executed.

4 Implementation on GPU

Fig.2 shows the process of feature matching on vertex and fragment program. After feature detection, binarization and invert operations are processed. Edge pixels of two images are black pixels with position (x, y). The detected edge images are stored in texture memory. From the texture memory, reference edge image is given to input of vertex program and target edge image is used to input of fragment program. We transform position (x, y) to texture coordinates (u, v) using mapping function in vertex program. In fragment program, texture re-sampling is executed at the coordinates (u, v) of target edge image. If the value of coordinates (u, v) is 0, the pixel of target edge image also regarded as edge pixel. That is, black pixels in new texture image represent correspondence



In Vertex Program

1. Define edge pixel position (x, y, 0) of Reference edge image
2. Calculate texture coordinate(u, v) using position(x, y) & changing parameters
   u = (x * scaling*rotation + translationX)/ width
   v = (y * scaling *rotation+ translationY)/ height

In Fragment Program

1. Texture re-sampling at (u, v) in Target edge image
2. Return of I(u,v) at position (x, y, 0)
   (I(u, v) : intensity of (u,v) coordination)
   if I(u,v) = 0, match
   otherwise, mismatch

Fig. 2. Feature matching in vertex & fragment program

of between reference edge pixel and target edge pixel. The values are written in new texture image and it is displayed. The number of black pixels means the matching cost of previously selected mapping parameters. Such processes are performed repeatedly in the possible range of parameters and calculate the value of matching cost as sum of overall black pixels in generated texture image. The maximum value of calculated matching cost used transformation parameter of between reference image and target image.

5 Experiments

To evaluate the performance of the image registration on GPU, each registration processing time in CPU and GPU are estimated. This test was executed in an environment using Intel Core2 CPU 6600 2.4GHz and the NVIDIA Geforce 8600 GTS GPU. As a standard in the speed measurement, only the image registration was calculated excluding initialization and configuration of OpenGL and Cg. Fig.3 shows the test image pairs:general aerial image pair [16] and multisensor image pair [17]. Left column images are reference images and Right column images are target images.

Table 1 show the estimated parameters and the registration speed of CPU and GPU. The major differences of the performance are the number of detected feature points and image size. If images contain many textures and image size is large, the necessary time for the image registration is much needed. The difference in the performance of between CPU and GPU is happened because of GPU architecture, internal parallelism, several vertex/fragment processors.

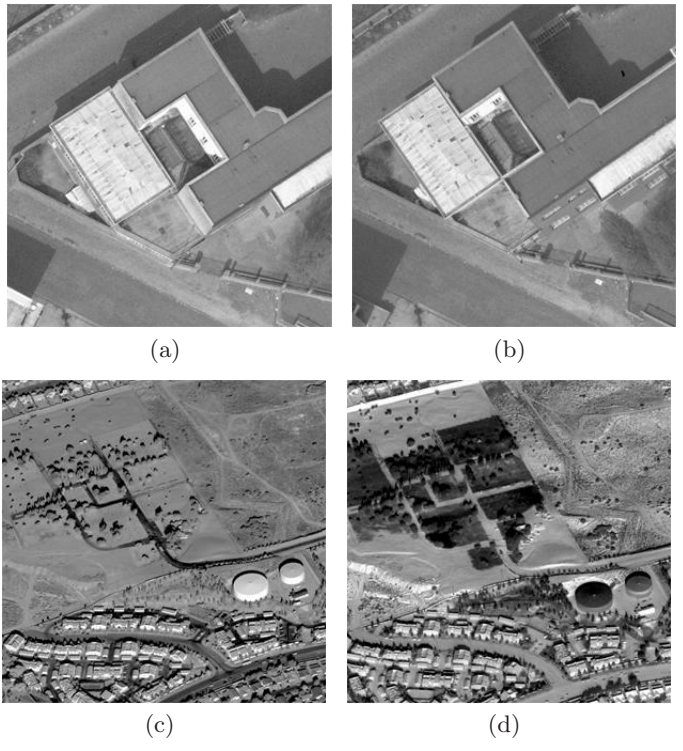


Fig. 3. (a)-(b)Aerial images (c)-(d)Multisensor images

Table 1. Estimated parameter and speed comparison in CPU/GPU

	Translation - x	Translation - y	Scale	Rotation (degree)	CPU (sec)	GPU (sec)
Image Pair -1 (600x600)	37.13	13.51	0.91	2.37	1.32	0.53
Image Pair -2 (1800x1800)	20.44	68.05	1.09	1.89	4.75	2.36

6 Conclusion

GPUs are likely the most cost-effective, high performance, processors available today. In recent, many researchers are interested in mapping general purpose computation(GPGPU) to graphics hardware. However, graphics hardware remains difficult to apply to non-graphics tasks for unusual programming model and various constraints of GPU. We implemented clustering-based image registration method on GPU using only transformation of texture coordinations in vertex program and re-sampling in fragment program. As a result, an improvement of over 50% in the execution speed compared to the speed when using

the CPU has been demonstrated through parallelism by GPU architecture and programming model.

References

1. Goshtasby, A.A.: 2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications. Wiley, Chichester (2005)
2. Barbara, Z., Jan, F.: Image Registration Methods: a survey. Image and Vision Computing 21(11), 977–1000 (2003)
3. Audette, M.A., Ferrie, F.P., Peters, T.M.: An Algorithmic Overview of Surface Registration Techniques for Medical Imaging. Medical Image Analysis 4, 201–217 (2000)
4. Bentoutou, Y., Taleb, N., Kpalma, K., Ronsin, J.: An Automatic Image Registration for Applications in Remote Sensing. IEEE Transactions on Geoscience and Remote Sensing 43(9) (2005)
5. Stockman, S.K., Benett, S.: Matching Images to Models for Registration and Object Detection via Clustering. IEEE Trans. Pattern Analysis Math. Intell. PAMI 4(3), 229–241 (1992)
6. Xiong, Y., Quek, F.: Automatic Aerial Image Registration Without Correspondence. In: IEEE International Conference on Computer Vision Systems (ICVS), pp. 25–32 (2006)
7. John, D.O., Luebke, D.: A Survey of General-Purpose Computation on Graphics Hardware. In: Eurographics 2005, State of the Art Reports, pp. 21–51 (August 2005)
8. Wong, T.T., Leung, C.S., Heng, P.A., Wang, J.: Discrete Wavelet Transform on Consumer-Level Graphics Hardware. IEEE Transactions on Multimedia 9, 668–673 (2007)
9. Matt, P., Randima, F.: GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation. Addison Wesley, Reading
10. McCormick, P.S., Inman, J., Ahrens, J.P., Hansen, C., Roth, G.: Scout: A Hardware-accelerated System for Quantitatively Driven Visualization and Analysis. In: Proceedings of IEEE Visualization, pp. 171–178 (2004)
11. Cornwall, J., Kelly, P.: Efficient Multiple Pass, Multiple Output Algorithms on the GPU. In: CVMP (2005)
12. Cornwall, J.L.: Efficient Multiple Pass, Multiple Output Algorithms on the GPU. In: 2nd European Conference on Visual Media Production, pp. 253–262 (2005)
13. Fung, J., Mann, S.: Computer Vision Signal Processing on Graphics Processing Units. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (2004)
14. Randima, F., Mark, J.: Kilgard The Cg Tutorial: The Definitive Guide to Programmable Real-Time Graphics. Addison Wesley, Reading
15. Canny, J.: A Computational Approach To Edge Detection. IEEE Trans. Pattern Analysis and Machine Intelligence 8, 679–714 (1986)
16. Visual Geometry Group, `http://www.robots.ox.ac.uk/~vgg`
17. Recon/Optical Inc., `http://www.reconoptical.com`

Prediction of Aortic Diameter Values in Healthy Turkish Infants, Children and Adolescents Via Adaptive Network Based Fuzzy Inference System

Bayram Akdemir[1], Salih Güneş[1], and Bülent Oran[2]

[1] Department of Electrical and Electronics Engineering, Selcuk University, 42075 Konya, Turkey
[2] Faculty of Medicine, Department of Pediatric Cardiology, Selcuk University, 42080 Konya, Turkey
{bayakdemir,sgunes}@selcuk.edu.tr, bulentoran@selcuk.edu.tr

Abstract. The aorta diameter size one of the cardiac value is very important to guess for child before adult age, due to growing up body. In conventional method, the experts use curve charts to decide whether their measured aortic diameter size is normal or not. Our proposed method presents a valid virtual aortic diameter result related to age, weight and sex. The proposed method comprises of two stages: (i) data normalization using a normalization method called Line Base Normalization Method (LBNM) that is firstly proposed by us, (ii) normalized aortic diameter prediction using Adaptive Network Based Fuzzy Inference Systems (ANFIS). Data set includes real Turkish infants, children and adolescents values and divided into two groups as 50% training -50% testing split of whole dataset to show performance of ANFIS. LBNM compared to three normalization methods including Min-Max normalization, Z-score, and decimal scaling methods. The results were compared to real aortic diameters values by expert with nine year experiences in medical area.

Keywords: Aortic Diameter, ANFIS, Line Base Normalization Method, Prediction.

1 Introduction

Echocardiography is a very important diagnostic imaging modality in the clinical practice of pediatric cardiology. Although it has been extensively used as a diagnostic tool, few studies are available in the literature that establishes echocardiographic reference values for the normal infants, children and adolescents population [1]. Thus, the echocardiographic measurements currently used are often based on values derived from small, non-randomized samples. Before the adult age for human, a body grows up every time and has more changes according to the previous time. Sometimes, if any surgical operation was in need especially for a child or an infant, experts may not decide for surgical operation due to growing body. Aortic diameter is one of the important parameter to must be known due to growing body size, weight and sex. Due to

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 498–505, 2008.
© Springer-Verlag Berlin Heidelberg 2008

blood pressure of heart output, aorta vessel expands and end of the beat become normal diameter. If aorta size is too big, due to blood pressure aorta vessel expands extremely. This expanding in the aorta vessel makes heart tire and may lead turning the blood back into heart due to this over flexibility. If an expert knows aortic diameter size is normal value for the subject than make decision if the measured diameter of the aorta is abnormal or not. In the conventional method, there are curve charts which are prepared previously hardcopy for experts. Any expert scrutinizes these charts to generate for solution if aortic diameter size measured by her/his is normal or not. In addition to these, curve charts may not include the current subject value.

As far as we know, this paper is the first study related to prediction of the aortic diameter size for healthy human before adult age in literature. In this study, a new normalization method called LBNM and ANFIS used together to help to physicians on what normal aortic diameter value is. For training and testing of ANFIS, real data sets were used. Features of aortic diameter dataset were explained in the section of material. There are some normalization methods for data normalization before prediction process. Among these, especially min-max normalization method is very used common for datasets which have different features. For example, one of the columns of dataset is weight and another column is related to year. Z-Score and decimal scaling methods are the other normalization method to normalize the dataset. After min-max normalization, the values of dataset are in the range of [0,1], but z-score and decimal scaling are scaled data in the range of varying from -1 to 1. Z score uses standard deviation to normalize data and centre of the standard deviation of the column equals to zero. Nearly all normalization methods generate new results through the column but our proposed method called Line Based Normalization Method (LBNM) deal with rows of dataset. After LBNM method, ANFIS used to predict normalized aortic diameter dataset belonging to Turkish healthy infant, children and adolescents. Also, this study helps expert to know what aortic diameter size is for current subject related to his/her weight, sex and age.

2 Material

Aortic diameters data were obtained from 2733 healthy infants and children aged from one day to 18 years in a single center in Middle Anatolia. In this study, our aim was to obtain normal aortic diameter values in a correct sample of normal Turkish infants, children and adolescents. The present study is conducted in Pediatric Cardiology Unit from January-1994 to January-2007. Healthy newborns, infants and children without cardiac heart disease or a history of cardiac involvement in infectious, hematological, neuromuscular, or metabolic disorders were included in this study. Most were outpatients referred for evaluation of a heart murmur, which was found to be innocent on clinical, radiological and electrocardiographic grounds. The aortic diameter was recorded by M-mode echocardiography at level of above the aortic valve. Internal aortic diameters were measured by means of a caliper in systole and diastole as the distance between the trailing edge of the anterior aortic wall and the leading edge of the posterior aortic wall (Figure 1).

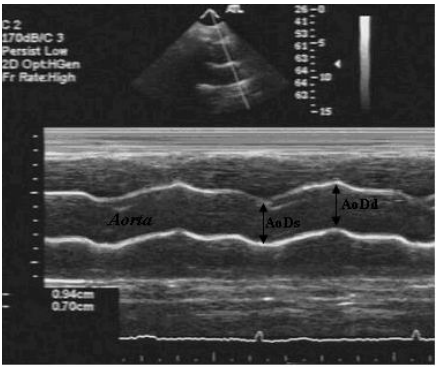


Fig. 1. Normal aortic diameter during cardiac cycle (systole and diastole) on M-mode

If the children were too restless, sedation was used (chloral hydrate, 50 mg/kg, per oral) before echocardiography. The neonates, infants, children, and adolescents taking part in this study represented a homogeneous sample of the normal healthy population. Their weights were all within the normal range on standard growth charts. They were examined by one of the two pediatric cardiologists to ensure that they had normal hearts before the echocardiograms were obtained. In all cases their chest roentgenogram and electrocardiogram recordings were within age appropriate normal limits.

All the subjects had complete cross sectional 2-D, and Doppler examinations. The children were examined in the supine position with the right shoulder slightly raised, during aortic diameter investigation.

3 The Proposed Method

The proposed method consists of two stages. In the first stage, the aortic diameter dataset is normalized in the range of [0, 1] using LBNM. In the second stage, the normalized aortic diameter values are given as input of ANFIS. In this study, aortic

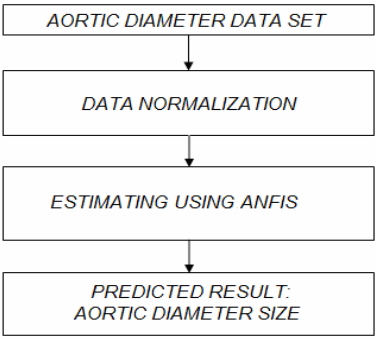


Fig. 2. General block diagram he block diagram of used method

diameter values dataset divided into two groups as 50% training-50% testing of whole dataset. The proposed method generates a possible aortic diameter value related to age, body weight, and sex of healthy subjects. In addition to LBNM, three normalization methods have been used and compared these normalization methods to each other. The flowchart of the proposed method is shown in Figure 2.

3.1 Line Based Normalization Method (LNNM)

All the attributes in any dataset may not always have a linear distribution among classes. If the non-linear classifier system is not used, data scaling or cleaning methods are needed both to transform data from original format to another space to improve the classification performance in pattern recognition applications. In this study, we have proposed a new data pre-processing LBNM in pattern recognition and medical decision making systems. In this method, the proposed data scaling method consists of two steps. In the first step, we have weighted data using following equation (1). In the second step, weighted data is normalized in the range of [0,1]. By this way, data is scaled in the basis of features used in dataset. Figure 3 shows the pseudo code of Attribute Based Data Normalization.

Input: d matrix with n row and m column
Output: weighted d matrix though column based data weighted method
1. Data is weighted by means of following equation.
for i=1 to n (such that n is the number of row in d matrix)
for j=1 to m (such that m is the number of features (attributes) in d matrix)

$$D_column = \frac{d_{i,j}}{\sqrt{(d_{i,1})^2 + (d_{i,2})^2 + ... + (d_{i,j})^2}} \quad (1)$$

end
end
2. Apply data normalization process after weighted process to 'D_column' matrix.

Fig. 3. The pseudo code of LBNM

3.2 Adaptive Network Based Fuzzy Inference System (ANFIS)

ANFIS is a class of adaptive networks that are functionally equivalent to fuzzy inference systems. To present the ANFIS architecture, this is a fuzzy Sugeno model in the framework of adaptive systems to facilitate learning and adaptation; fuzzy if-then rules based on a first order Sugeno model are considered [2,3]:

Rule 1: If (x is A_1) and (y is B_1) then ($f1 = p_1x + q_1y + r_1$)
Rule 2: If (x is A_2) and (y is B_2) then ($f2 = p_2x + q_2y + r_2$)

Figure 4 presents the ANFIS architecture.

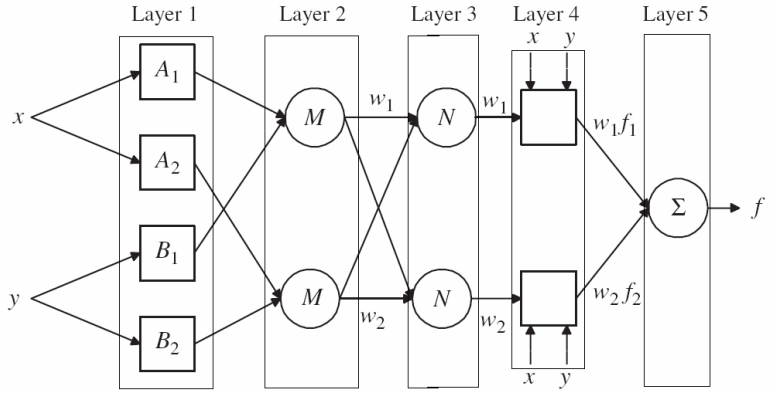


Fig. 4. The architecture of ANFIS

Layer 1: Every node *i* in this layer is an adaptive node with a node function

$$O_i^1 = \mu_{A_i}(x), i = 1, 2, \tag{1}$$

$$O_i^1 = \mu_{B_{i-2}}(y), i = 3, 4, \tag{2}$$

where x and y are the inputs to node i and A_i (or B_{i-2}) is a linguistic label related to this node. In other words, $O_{1,i}$ is the membership degree of a fuzzy set A (=A_1, A_2, B_1, B_2) and it explains the degree to which the given input x (or y) satisfies the quantifier A. For instance, if the bell shaped membership function is employed; $\mu_{Ai}(x)$ is shown as follows:

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \tag{3}$$

where $\{a_i, b_i, c_i\}$ are the parameter set. As the values of these parameters change, the bell-shaped function varies accordingly.

Layer 2: Every node in this layer is a fixed node labeled Π, whose output is the product of all the incoming signals:

$$O_i^2 = w_i = \mu_{A_i}(x)\mu_{B_i}(y), i = 1, 2 \tag{4}$$

Each node output shows the firing strength of a rule.

Layer 3: Every node in this layer is a fixed node labeled N. The *ith* node calculates the ratio of the *ith* rule's firing strength to the sum of all rules' firing strengths:

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2}, i = 1, 2 \tag{5}$$

For convenience, the outputs of this layer are called normalized firing strengths.

Layer 4: Every node *i* in this layer is an adaptive node with a node function

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i), i = 1,2 \tag{6}$$

where $\overline{w}_i$ is a normalized firing strength from layer 3 and $\{ p_i, q_i, r_i \}$ is the parameter set of this node. These parameters are referred to as consequent parameters.

Layer 5: The single node in this layer is a fixed node labeled $\sum$, which computes the overall output as the summation of all incoming signals:

$$O_i^5 = \sum_{i=1}^{2} \overline{w}_i f_i = \frac{\sum_{i=1}^{2} w_i f_i}{w_1 + w_2} \tag{7}$$

In order to perform fuzzy inference system in ANFIS, Training FIS optimization method was set subtractive clustering. Then subtractive clustering is partitioned the input data according to dimension of dataset and automatically tuned the input-output membership functions, Hybrid learning algorithm is used in order to identify the optimal values of optimization parameters including consequent parameters and premise parameters. The used parameters for ANFIS are 0.5 (Range of influence), 1.25 (Squash factor), 0.5 (Accept ratio), 0.15 (Reject ratio) and for three inputs created two memberships functions. ANFIS trained using 1000 epoch. [2,3,4].

4 Result and Discussion

Aortic diameter size directly related to heart healthy is very important to guess. Our proposed method helps to experts guess what possible aortic diameter size that belongs to normal healthy person must be. The expert gives final decision whether the subject healthy is normal or not according to his/her own idea.

Our proposed method comprises of two stages: (i) data normalization using a normalization method called Line Base Normalization Method (LBNM) that is firstly proposed by us, (ii) normalized aortic diameter prediction using Adaptive Network Based Fuzzy Inference Systems (ANFIS). Data set includes real Turkish infants, children and adolescents values and divided into two groups as 50% training -50% testing split of whole dataset to show performance of ANFIS. LBNM compared to three normalization methods including Min-Max normalization, Z-score, and decimal scaling methods. The obtained mean square error (MSE) results from ANFIS combined with normalization methods Min-Max, Decimal scaling, Z-score and LBNM are 0.004143, 0.000575, 0.004457, and 0.000097, respectively. In order to evaluate the proposed method, we have used the mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), T score, R^2 value, and Covariance [5-8]. While Table 1 shows the training results of ANFIS, Table 2 presents the testing results of ANFIS on the prediction of aortic diameter value.

Table 1. Averaged training results obtained from ANFIS with raw and normalization methods included MAE, MSE, RMSE, T, R, Covariance and AD using two-fold cross validation

	MAE	MSE	RMSE	T	R^2	Cov
Original	1.74124	5.54062	2.3524	0.78078	0.9854	12.484
Min-Max	0.04528	0.003791	0.06155	0.79641	0.9849	12.764
Decimal	0.01693	0.000528	0.02298	0.79095	0.9861	12.198
Z-Score	0.04543	0.003827	0.061.86	0.79323	0.9847	12.827
LBNM	**0.0045**	**20.95E-6**	**0.007.837**	**0.99747**	**0.9998**	**1.1712**

Table 2. Averaged testing results obtained from ANFIS with raw and normalization methods included MAE, MSE, RMSE, T, R, Covariance and AD using two-fold cross validation

	MAE	MSE	RMSE	T	R^2	Cov
Original	1.78259	5.97863	2.44201	0.7630	-434.64	12.960
Min-Max	0.04667	0.004143	0.06434	0.7774	-10.794	13.341
Decimal	0.01759	0.000575	0.02399	0.7735	-3.1966	12.735
Z-Score	0.04729	0.004457	0.06676	0.7598	-11.691	13.844
LBNM	**0.00488**	**0.000097**	**0.009**	**0.9960**	**0.7967**	**1.4955**

As can be seen from these results, the best method is combination of LNBM and ANFIS on the estimation of aortic diameter value. Also, this method can be useful for physicians on the estimation of aortic diameter value based on subject's age, weight, and sex.

5 Conclusion

In this study, we have proposed a novel normalization method called line based normalization method (LBNM) and also applied to prediction of aortic diameter value based on the information of subject's age, weight, and sex. LBNM combined with ANFIS and used as data pre-processing. Also, this normalization method has been compared to other normalization methods including min-max normalization, decimal scaling, and z-score. The advantage of this study is that the proposed method can be used instead of chart method in medical area in the prediction of aortic diameter value. Conventionally, experts use chart, prepared through the long years, to make decision. Our proposed method encourage us to advice the experts to give up the charts.

Acknowledgments. This study has been supported by Scientific Research Project of Selcuk University.

References

1. Feigenbaum, H.: Echocardiography, pp. 658–675. Lea&Febiger, Philadelphia (1994)
2. Güler, İ., Übeyli, D.E.: Adaptive Neuro-fuzzy Inference System for Classification of EEG Signals Using Wavelet Coefficients. Journal of Neuroscience Methods 148, 113–121 (2005)
3. Jang, J.S.R.: Self-learning Fuzzy Controllers Based on Temporal Back Propagation. IEEE Trans. Neural Network 3(5), 714–723 (1992)
4. Polat, K., Güne, S.: Ahybrid Medical Decision Making System Based on Principles Component Analysis, k-NN Based Weighted Pre-processing and Adaptive Neuro-fuzzy Inference system. Digital Signal Processing 16(6), 913–921 (2006)
5. Razmi-Rad, E., Ghanbarzadeh, B., Mousavi, S.M., Emam-Djomeh, Z., Khazaei, J.: Prediction of Rheological Properties of Iranian Bread Dough from Chemical Composition of Wheat Flour by Using Artificial Neural Networks. J. Food Eng. 81(4), 728–734 (2007)
6. Kaveh, N., Sh., J., Ashrafizadeh, S.N., Mohammadi, F.: Development of An Artificial Neural Network Model for Prediction of Cell Voltage and Current Efficiency in a Chlor-alkali Membrane Cell. Chemical Engineering Research and Design (Article in press)
7. Esena, H., Inallib, M., Sengurc, A., Esena, M.: Predicting Performance of a Ground-source Heat Pump System Using Fuzzy Weighted Pre-processing-based ANFIS. Building and Environment (Article in press)
8. Betchler, H., Browne, M.W., Bansal, P.K., Kecman, V.: New Approach to Dynamic Modelling of Vapour Compression Liquid Chillers: Artificial Neural Networks. Applied Thermal Engineering 21(9), 941–953 (2001)

Feature Extraction and Classification for Graphical Representations of Data

Jinjia Wang[1,2], Jing Li[3], and Wenxue Hong[2]

[1] College of Information Science and Engineer, Yanshan University, Qinhuangdao 066004
[2] College of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China
[3] Colleges of Science, Yanshan University, Qinhuangdao 066004, China
wjj@ysu.edu.cn, hongwx@ysu.edu.cn

Abstract. The barycentre graphical feature extraction method of the star plot is proposed, based on the graphical representation of multi-dimensional data. Because for the different feature order the same multi-dimensional data lead to the different star plots, and extract the different barycentre graphical features, which affect the classification error of the classifiers. The novel feature order method based on the improved genetic algorithm (GA) is proposed. Meanwhile the traditional feature order method based on the feature selection is researched and the traditional vector feature extraction methods are researched. The experiments results of the 4 real data sets show the classification effectiveness of the new graphical representation and graphical features.

Keywords: Star plot, graphical representation, graphical features extraction, feature order.

1 Introduction

The issue of representation is an essential aspect of pattern recognition and is different from classification [1]. It largely influences the success of the stages to come. This is a promising direction. Building proper representations has become an important issue in pattern recognition [2]. For a long time this idea has been restricted to the reduction of overly large feature sets to the sizes for which generalization procedures can produce significant results, given the cardinality of the training set. We may learn the class of objects by studying the domain of their corresponding representations. Several methods have been studied based on feature selection as well as linear and kernel feature extraction [3]. Very advanced methods are needed to find such 'hidden' subspaces in which classes are well separable [4].

The one keystone in pattern recognition is that how the mathematic features are selected and extracted by the learning observations [2]. Although one also often utilize the physical and structural features to recognize the object, as these features are easily found out by human. But the capacity of extracting the mathematic features using computer, such as statistical mean, correlation, eigenvalue and eigenvector of observation covariance, is more superior to by human. These two sides are not integrated. We proposed a method implementing this integration.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 506–513, 2008.
© Springer-Verlag Berlin Heidelberg 2008

As we know, the graphical representation or graphical analysis for multi-dimensional data in multivariate analysis is a very useful tool [5]. One commonly used graphical representation form is the 'star plot', in which the profile lines are placed on spokes so that the profile plot looks a bit like a star. Each dimension is represented by a line segment radiating from a central point. The ends of the line segment are joined. The length of the line segment indicates the value of the corresponding dimension. So the star plot is a very useful way of displaying the data with fairly large number of variables. But the graphical representation for multi-dimensional data has a limited application in pattern recognition. In pattern recognition, the data set typically (but not necessarily) is a subset of R^d, if the dimension of input feature space is d. In general, the normalized vector data is trained and classified. Based on the representation question, we propose a new data representation method in pattern recognition: graphical representation. Based on the graphical representation of data, new feature extraction method from the graphical representation (such as the star plot) of data is proposed. From the star plot of a multivariate data, we see the irregular polygonal shape by encircled the variance value on spokes. Based on the shape, we propose the area features and barycentre features for each observation, the number of which is both the same as the dimension of the observation. These new graphical features extend the basic feature concept. Moreover, these new graphical features establish an integration of the geometrical feature and mathematic feature. The experimental result of our previous primary work [6] indicates the validity of the new graphical feature for the K nearest neighbor classifier. But the proposed new graphical feature is affected from the feature order. That paper does not deal with the important question. The different feature order lead to the different graphical feature and the different graphical feature lead to the different classification error. So this paper thought over and solved the feature order question based on the proposed feature order methods. This is our contribution. The graphical features of the optimal feature order are evaluated by the popular classifiers such as the traditional classifier [2], which is compared with the other relational methods. The evaluation of the consequent classifier is done through leave one out cross validation (LOOCV) procedure repeated ten times. Experiments with several standard benchmark data sets show the effectiveness of the new graphical features and the feature order. One hand, the classification errors of barycentre feature under the original feature order of multivariate data are better or worse than the traditional feature extraction method. But the classification errors of barycentre feature under the feature order based on the improved genetic algorithm (GA) are better than the traditional feature extraction methods. On the other hand, the classification errors of barycentre feature under the feature order based on the improved genetic algorithm are better than the traditional feature order method based on the feature selection.

2 Graphical Representation and Graphical Features

The star plot is a simple means of multivariate visualization, which represents the value of an attribute through the length of lines radiating from the icon's center. This method provides an overall impression of change of variable values across subjects. As to the vector data, we should not be limited in the only data graphical representation, but

should full utilize data graphical analysis. That is, we should look for a method to mining the graphical features of the star plot. So we propose the barycentre graphical features of star plot. We first rescale each variable to range from 0 to 1. For the rescaled observation ($r_1, r_2 \cdots, r_d$), its star plot include d triangle, which is a visional shape feature. Each triangle has a barycentre point, and a whole star plot has d barycentre with d amplitude value and d angle value. We neglect the angle value. So the original data are changed to the barycentre features with the same size. The barycentre graphical amplitude features of the observation absi can be calculated as the following equation. The angle wi= 2π/d is the angle between two conterminous dimension. This detail refers to the paper [6].

$$abs_i = f(r_i, r_{i+1}, \omega_i) = \frac{\sqrt{r_i^2 + r_{i+1}^2 + 2r_i r_{i+1} \cos \omega_i}}{3}, i = 1,2,\cdots, d \tag{1}$$

3 Graphical Representation Approach under the Feature Order

3.1 Experiments of the Iris Barycentre of under the Exhaustive Feature Order

The mission of feature order is that selects all of features which quantity is d from a group of features which quantity is d, and requires different feature permutation order. The matter of feature order isn't discussed generally in the existing pattern recognition, as the feature order has no influence on every current classifiers method. So the method of feature order is very little. The graphic feature of star plot is affected by the feature order. From the experimental results, not only feature order has affect on the graphic feature, but also has effect on the performance of the classifiers commonly used. Such as for the 4 dimension data [4.7 10.3 2.1 8.6], one feature order may be [8.6 4.7 2.1 10.3]. The star plot of them is shown as Fig.1. We observe the different feature order leads to the different star plot for the same data. This may lead to the different graphical feature and the different classification performance.

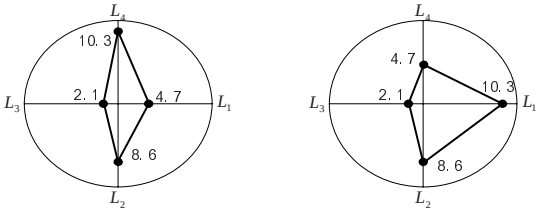


Fig. 1. The star plot of the same data under different feature order

In this work several statistical classifiers were used, because of their simplicity and robustness. There approaches include: (1) Parametric approach: LDA; (2) Non-parametric approach: K-nearest neighbors (KNN) with K=1 and K=3, and the Euclidean distance function. (3) Support vector machines (SVM): SVM are a kind of learning machine based on statistical learning theory.

To choose d features from features d and require different feature permutation order, all of possible combinatorial numbers is $q = d!$. For IRIS data, if $d = 4$, $q = 4\times3\times2\times1 = 24$. If $d = 20$, $q = 20$. This method is named exhaustive method. The 24 feature order of the exhaustive method of Iris data is exercised below. We studied the experimental results of the barycentre graphical features of the Iris data under the exhaustive feature order using 1NN and LOOCV. We find the error of 1NN of the traditional normalization feature for all feature order is the same. This note the traditional feature is not sensitive for the feature order. But for the graphical feature, the best classification performance 0.0267 is achieved using the 8 feature order. The rate is 8/24. This note the best feature order is easy to find. So the 3 independent feature orders for the Iris data have to consider, which is shown as [1 2 3 4], [1 2 4 3] and [1 3 2 4], other feature order can be neglected.

3.2 The Feature Order Method

From former analysis, for high dimension data, the calculation amount for searching a optimal feature order is too large to realize. Using a feasible algorithm is absolutely of necessity. Only exhaustive method can guarantee optimal results we get. All algorithms are still exhaustive method principally, nothing else than the calculation amount is decreased because of adopting some searching technology.

First the feature order method of feature selection is given as follows. Feature order is generally involved in feature selection, and feature selection is a classical matter of pattern recognition [7]. The first feature order method is that, the classification criterion value is calculated using the single feature according to some rule, and then order these features according to the size of the criterion value. If the certain feature order under some rule is an optimal feature order, the feature order of the graphic feature of star plot is very simple. The classification criterions J commonly used include: classification criterion based on Euclidean distance, classification criterion based on probability distribution (such as Euclidean distance, Mahalanobis distance), classification criterion based on LOOCV performance evaluation of 1NN. Note the criterion values of different methods don't have comparability. So we can get one feature order: $J(x_1) > J(x_2) > J(x_3) > \ldots > J(x_{d-1}) > J(x_d)$.

Obviously, these feature order method only gives a kind of feature order, which may be not suitable for our question. The feature order should is associated with the data and classifier. So we proposed the improved GA to order the feature.

GA is randomized search and optimization techniques guided by the principles of evolution and natural genetics. They are efficient, adaptive and robust search processes, producing near optimal solutions and have a large amount of implicit parallelism. The utility of GA in solving problems that are large, multi-modal and highly complex has been demonstrated in several areas. Good solutions are selected and manipulated to achieve new and possibly better solutions. The manipulation is done by the genetic operators (selection, mutation, and crossover) that work on the chromosomes in which the parameters of possible solutions are encoded. In each generation of the GA, the new solutions replace the solutions in the population that are selected for deletion.

We consider integral coded GA for the networks parameters. Each individual is composed of the integer from 1 to the maximum dimension d. The length of individual

is d and each gene must be not same. The individuals are created randomly by the Matlab function order=randperm(d), which realized one order of the d dimension. For example, the original feature order is [1 2 3 4] for the iris data, the initial feature could be [2 3 1 4].

The genetic operation includes selection operation, crossover operation and mutation operation. The selection operation mimics the 'survival of the fittest' concept of natural genetic systems. Here the individuals are selected from a population to create a mating pool. The probability of selection of a particular individual is directly or inversely proportional to the fitness value depending on whether the problem is that of maximization or minimization. In this work, the individual with a larger fitness value, i.e. better solution to the problem, receive correspondingly larger numbers of copies in the mating pool. The size of the mating pool is taken to be same as that of population. The crossover operation is a probabilistic process that exchanges information between two parent individual and generates two offspring for the next population. Here one-point crossover with a fixed crossover probability of pc=1 is used. The mutation operation used the uniform mutation. Each individual undergoes uniform mutation with a fixed probability pm =0.005. Elitism is an effective means of saving early solutions by ensuring the survival the fittest individual in each generation. The elitism puts the best individual of old generation into the new generation.

A fitness function value is computed for each individual in the population, and the objective is to find the individual that has the highest fitness for the problem considered. The fitness function is the classification correct rate of the three classifiers (LDC, 1NN, 3NN and SVM) with the LOOCV.

As the stopping rule of maximum generation or the minimum criterion of relative change of the fitness values is satisfied, the GA process stops and the solution with the highest fitness value is regarded as the best feature order.

4 Experiments and Results

Several standard benchmark corpora from the UCI Repository of Machine Learning Databases and Domain Theories (UCI) have been used [8]. A short description of these corpora is given below. 1) Iris data: This data set consists of 4 measurements made on each of 150 iris plants of 3 species. The two species are iris setosa, iris versicolor and iris virginica. 2) Liver data: This data set consists of 6 measurements made on each of 345 data of 2 classes. 3) Wisconsin breast cancer data: This data set consists of 9 measurements made on each of 683 data (after removing missing values) of 2 classes (malignant or benign). 4) Sonar data: This data set consists of 60 frequency measurements made on each of 208 data of 2 classes ("mines" and "rocks").

We use the barycentre graphical features under the different order which include the original order, the order use the between and within-class scatter distance, the order use the Euclidean distance, the order use the Mahalanobis distance, the order use the sequential forward selection (SFS), the order use the sequential backward selection (SBS), the order use the sequential forward floating search method (FLS), and the order use the GA [9]. For the comparison, we also use the linear normalized feature of the range U[0 1], the feature with the principal component analysis (PCA) [10] with all the d dimensions, the feature with the kernel principal component analysis (KPCA) [10]with all the d dimensions using kernel parameter 4 and the C value 100.

Table 1. Classification error using LOOCV for Iris dataset

feature \ classifier	LDA	1NN	3NN	SVM
linear normalized feature U[0 1]	0.0200	0.0467	0.0467	0.0333
PCA feature	0.0200	0.0400	0.0400	0.0467
KPCA feature	0.0133	0.1067	0.1133	0.0467
barycentre features with original order	0.0200	0.0333	0.0267	0.0333
barycentre features with scatter distance order	0.0200	0.0333	0.0267	0.0333
barycentre features with Euclidean distance order	0.0400	0.0533	0.0460	0.0467
barycentre features with Mahalanobis distance order	0.0333	0.0667	0.0800	0.1467
barycentre features with SFS order	0.0333	0.0467	0.0533	0.0400
barycentre features with SBS order	0.0400	0.0467	0.0533	0.0333
barycentre features with FLS order	0.0200	0.0333	0.0267	0.0533
barycentre features with GA order	0.0200	0.0333	0.0267	0.0333

Table 2. Classification error using LOOCV for Bupa Liver-disorders dataset

feature \ classifier	LDA	1NN	3NN	SVM
linear normalized feature U[0 1]	0.3072	0.3739	0.3652	0.2696
PCA feature	0.3072	0.3768	0.3623	0.4029
KPCA feature	0.3855	0.4609	0.3855	0.3855
barycentre features with original order	0.3449	0.4058	0.4000	0.3101
barycentre features with scatter distance order	0.3391	0.3507	0.3797	0.3304
barycentre features with Euclidean distance order	0.3304	0.4145	0.4000	0.4087
barycentre features with Mahalanobis distance order	0.3188	0.4377	0.4290	0.4232
barycentre features with SFS order	0.3304	0.4377	0.4087	0.3913
barycentre features with SBS order	0.3246	0.4493	0.4319	0.4058
barycentre features with FLS order	0.4000	0.4551	0.4667	0.4174
barycentre features with GA order	0.2812	0.3507	0.3275	0.2957

The setting parameter is as the following. The population is 20, the maximum generation is 50, the selection operators is the roulette wheel , the crossover operators is the single crossover with the crossover rate 1, the mutation operators is the uniform mutation with the mutation rate 0.02.

For the 1NN, 3NN and LDC classifier we use PRTOOLS toolbox [12]. For the SVM classifier with radial basis kernels with kernel scale value 0.5 and C value 1000, we used SVMLIB toolbox [13]. The whole system was made in MATLAB.

From Table 1-4, for the four classifiers, the classification performance of the barycentre features with GA order is superior to the PCA features, KPCA features, linear normalized feature, Gaussian normalized feature. From Table 1-4, for the four classifiers, the classification performance of the barycentre features with GA order is superior to the barycentre features with the other order. Once an optimal order is obtained, the train time and recognition time should decrease.

Table 3. Classification error using LOOCV for breast-cancer-Wisconsin dataset

feature \ classifier	LDA	1NN	3NN	SVM
linear normalized feature U[0 1]	0.0395	0.0454	0.0322	0.0454
PCA feature	0.0395	0.0425	0.0322	0.0747
KPCA feature	0.0395	0.0351	0.0395	0.1947
barycentre features with original order	0.0395	0.0439	0.0264	0.0293
barycentre features with scatter distance order	0.0395	0.0351	0.0293	0.0278
barycentre features with Euclidean distance order	0.0366	0.0351	0.0249	0.0293
barycentre features with FLS order	0.0395	0.0337	0.0337	0.0293
barycentre features with GA order	0.0366	0.0264	0.0190	0.0249

Table 4. Classification error using LOOCV for Sonar dataset

feature \ classifier	LDA	1NN	3NN	SVM
linear normalized feature U[0 1]	0.2452	0.1250	0.1635	0.0962
PCA feature	0.2452	0.1731	0.1827	0.1202
KPCA feature	0.1875	0.2837	0.3029	0.3365
barycentre features with original order	0.2404	0.1442	0.1346	0.1154
barycentre features with scatter distance order	0.2500	0.1490	0.1538	0.1058
barycentre features with Euclidean distance order	0.2548	0.1779	0.1971	0.1731
barycentre features with FLS order	0.2500	0.1058	0.1442	0.1827
barycentre features with GA order	0.1731	0.0865	0.0913	0.0673

5 Conclusion

Based on the concept of the graphical representation, this paper proposes the concept of the graphical features and gives barycentre graphical features based on star plot.

And the feature order is researched using the improved feature selection and improved Genetic Algorithms because the feature order affect the graphical features and affect the classification performance. The experiments results using four data sets proved our idea and methods. The results shows that the proposed graphical features with the GA order can achieve high classification accuracy.

To fully investigate the potential of the graphical features, more comprehensive experiments can be performed. One possible future direction is the new graphical features which make the class reparability. Another possible future direction is the new feature order methods.

Acknowledgments. This work was supported by National Natural Science Foundation of China (No60504035). The work was also partly supported by the Science Foundation of Yanshan University for the Excellent Ph.D Students.

References

1. Robert, P.W.D., Pekalska, E.: The Science of Pattern Recognition. Achievements and Perspectives, Studies in Computational Intelligence 63, 221–259 (2007)
2. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification and Scene Analysis, 2nd edn. John Wiley and Sons, New York (2000)
3. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical Pattern Recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(1), 4–37 (2000)
4. Pudil, P.J., Novovi´cova, K.J.: Floating Search Methods in Feature Selection. Pattern Recognition Letters 15(11), 1119–1125 (1994)
5. Gao, H.X.: Application Statistical Multianalysis. Beijing University Press, Beijing (2005)
6. Wang, J.J., Hong, W.X., Li, X.: The New Graphical Features of Star Plot for K Nearest Neighbor Classifier. In: Huang, D.-S., Heutte, L., Loog, M. (eds.) ICIC 2007. LNCS (LNAI), vol. 4682, pp. 926–933. Springer, Heidelberg (2007)
7. Christopher, M.B.: Pattern Recognition and Machine Learning. Springer, New York (2006)
8. Asuncion, A., Newman, D.J.: UCI Machine Learning Repository. University of California, School of Information and Computer Science, Irvine (2007), `http://www.ics.uci.edu/~mlearn/MLRepository.html`
9. Huang, C.L., Wang, C.J.: A GA-based Feature Selection and Parameters Optimization for Support Vector Machines. Expert Systems with Applications 31, 231–240 (2006)
10. Scholkopf, B., Smola, A.J., Muller, K.R.: Kernel Principal Component Analysis. In: Artificial Neural Networks-ICANN 1997, Berlin, pp. 583–588 (1997)
11. Duin, R.P.W., Juszczak, P., Paclik, P., Pekalska, E., Ridder, D.D., Tax, D.M.J.: PRTools4, A Matlab Toolbox for Pattern Recognition, Delft University of Technology (2004)
12. Chang, C.C., Lin, C.J.: LIBSVM: a Library for Support Vector Machines (2001), `http://www.csie.ntu.edu.tw/~cjlin/libsvm`

Improving Depth Resolution of Diffuse Optical Tomography with Intelligent Method

Hai-Jing Niu[1], Ping Guo[1, 2], and Tian-Zi Jiang[3]

[1] Image Proc. & Patt. Recog. Lab, Beijing Normal University, Beijing 100875, China
[2] School of Comp. Sci. & Tech, Beijing Institute of Technology, Beijing 100081 China
[3] National Lab. of Patt. Recog., Inst. of Auto., CAS, Beijing 100081, China
niuhjing@163.com, pguo@ieee.org, jiangtz@nlpr.ia.ac.cn

Abstract. Near-infrared diffuse optical tomography imaging (DOT) suffers from a poor depth resolution due to the depth sensitivity decreases markedly in tissues. In this paper, an intelligent method, which is called layered maximum-singular-values adjustment (LMA), is proposed to compensate the decrease of sensitivity in depth dimension, and hence obtain improved depth resolution of DOT imaging. Simulations are performed with a semi-infinite model, and the simulated results for objects located in different depths demonstrate that the LMA technique can improve significantly the depth resolution of reconstructed objects. The positional errors of less than 3 mm can be obtained in the depth dimension for all depths from -1 cm to -3 cm.

Keywords: Depth resolution, reconstruction, diffuse optical tomography, layered maximum singular values.

1 Introduction

Near-infrared diffuse optical tomography (DOT) imaging has been of increasing interest in breast cancer diagnosis and brain function detection in recent years [1-3]. It can quantify hemoglobin concentrations and blood oxygen saturation in tissue by imaging the interior optical coefficients of tissues [2, 4], and perform non-invasive detection and diagnosis.

A number of advantages can be found in near-infrared DOT imaging, including portability, real time imaging and low instrumentation cost [5, 6], but it is an undeniable fact that the depth resolution for DOT imaging has always been poor[7]. For instance, it cannot discriminate blood volume change of the cerebral cortex with brain activity and the skin blood volume change with autonomic nerve activity, which limits its clinical application. In order to improve the depth resolution of DOT imaging, Pogue *et al.* [8] show that using spatially variant regularization parameters can obtain the constant spatial resolution in imaging fields. Zhao *et al.* [9] point out that a sigmoid adjustment method in forward matrix can improve effectively the depth resolution of DOT imaging. For these two methods, there are several parameters in the adjustment formulas and they need users to choose, which are not easy because of no theoretical guidance in selecting them. In practice, users determine these parameters based on experience, which are sometimes subjective and inconvenient.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 514–520, 2008.
© Springer-Verlag Berlin Heidelberg 2008

In this paper, an intelligent method called layered maximum-singular-value adjustment (LMA) is proposed to improve the depth resolution of DOT imaging. The layered maximum-singular-value can reflect the exponential change in optical density from the superficial to deep layers. In order to decrease the sensitivity difference between superficial and deep layers, we multiply the elements in the sensitivity matrix with the layered maximum singular values matrix in inverse layered order. Simulations are performed using a semi-infinite model to study the proposed LMA method.

2 DOT Theory and Method

Photon propagation in tissue can be described by the diffusion equation [10, 11] when tissue acts as a highly scattering medium, which in the frequency domain is given by

$$-\nabla \cdot \kappa(\mathrm{r})\nabla\Phi(\mathrm{r},\omega)+\left(\mu_{\mathrm{a}}+\frac{\mathrm{i}\omega}{\mathrm{c}}\right)\Phi(\mathrm{r},\omega)=\mathrm{q}_{0}(\mathrm{r},\omega) \tag{1}$$

where $\Phi(\mathrm{r},\omega)$ is a complex number radiance at position r with modulation frequency ω, $q_0(\mathrm{r},\omega)$ is an isotropic source, c is the wave speed in the medium, μ_{a} is an absorption coefficient and $\kappa(\mathrm{r})$ is an optical diffusion coefficient. On the condition of semi-infinite assumption, the analytic solution of the diffusion equation can be obtained when considering only changes in the absorption portion. The solution can be written as in the following form [11]:

$$\Delta OD=-\ln\left(\frac{\Phi_{pert}}{\Phi_{0}}\right)=\int\Delta\mu_{a}(r)L(r)dr \tag{2}$$

where ΔOD is the change in optical density. Φ_0 is the simulated photon fluency in a semi-infinite model, and Φ_{pert} is the simulated photon fluence with the absorbers included. In this semi-infinite model, we set the imaging field area 6.0×6.0 cm^2 in x-y plane, and the depth for this model in z direction is 3 cm. We assume that the absorption changes will occur at depths of -1.0 to -3.0 cm. The background absorption and reduced scattering coefficients for the medium are 0.1 cm^{-1} and 10cm^{-1}, respectively. A hexagonal configuration of 7 source and 24 detectors [13] is arranged on the surface of the imaging region, which produce 168 measurements. Random Gauss-distributed noise was added to the simulated data to achieve measurements with approximate signal to noise ratio (SNR) of 1000.

This equation (2) is known as the modified Beer-Lambert law. And it can be generalized for a set of discrete voxels and be written as $\mathbf{y}=\mathbf{A}\mathbf{x}$, where $\mathbf{y}$ is the vector of measured changes in optical density from all the measurements, $\mathbf{x}$ is the vectors of the unknown change in absorption coefficient in all of the voxels, and matrix $\mathbf{A}$ describes the sensitivity that each measurement has to the change in absorption within each voxel in the medium.

For the inverse problem, due to few measurements than unknown optical properties for all voxels, the linear problem is ill-posed and it is inversed with Tikhonov regularization regulation [14]:

$$\mathbf{x} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T + \lambda \mathbf{I})^{-1} \mathbf{y}, \tag{3}$$

The coefficient λ in the equation is a regularization parameter, and it holds constant value 0.001 during the whole image reconstruction. The voxel size in the reconstructed images is $1\times1\times1$ mm^3.

The sum of matrix A by rows indicates the overall sensitivity of each voxel from all measurements, and it displays an exponential decrease distribution in an increased depth dimension [9], which makes the deep object usually be reconstructed bias toward the upper layers [15].

3 LMA Method

Figures 1(a) and 1(b) show the sensitivity distribution in *x*-*z* plane and its sensitivity curve in *z* direction, respectively. This illustrates that the photon density drops off exponentially in the depth dimension. As a result, in this study, an intelligent method called layered maximum-singular-values adjustment (LMA) was developed to adjust the decrease of sensitivity, which would be expected to obtain improved depth resolution for the reconstructed images. More specific, this LMA intelligent method includes the following steps.

Firstly, we assume that the imaging field contains *L* layers. And then we calculate the maximum-singular-value for each layer in the forward matrix and they are represented with M_L, M_{L-1}, …, M_1 from the *L*th layer to the first layer, respectively.

Secondly, we assign the value M_L to be the adjustment coefficient for the first layer and M_{L-1} is the adjustment coefficient for the second layer, and so on, the M_1 is the adjustment coefficient for the bottom layer (i.e. the *L*th layer). That is, we utilize these singular values in inverse order. In this way, we obtain the adjusted forward matrix $\mathbf{A}^{\#}$, and it is written as follows:

$$\mathbf{A}^{\#} = \mathbf{A}\mathbf{M}, \quad \mathbf{M} = \begin{bmatrix} M_L & & & & & & \\ & \ddots & & & & & \\ & & M_L & & & & \\ & & & \ddots & & & \\ & & & & M_1 & & \\ & & & & & \ddots & \\ & & & & & & M_1 \end{bmatrix} \tag{4}$$

where $\mathbf{M}$ is a diagonal matrix. These layered singular values have the following properties: $M_1 > M_2 > \ldots\ldots > M_L$, which exhibits an exponential decrease configuration from the superficial to deep layers, as shown in Fig. 1. The exponential drop can be considered as an effect of the decrease of photon sensitivity with the increase of penetrated photon depth. Equal adjustment coefficients were assigned to every voxel within the same layer. In this way, we obtain its new reconstruction equation for this LMA: $\mathbf{x} = \mathbf{M}\mathbf{A}^T (\mathbf{A}\mathbf{M}^2\mathbf{A}^T + \lambda \mathbf{I})^{-1} \mathbf{y}$.

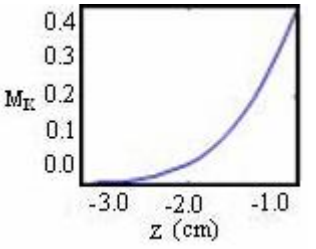


Fig. 1. M_k is the maximum singular values for layered sensitivity matrix and *k* is changed from 1 to *L*. The imaging depth is from -1cm to -3cm.

As for the adjusted forward matrix $\mathbf{A}^{\#}$, the sensitivity distribution in *x*-*z* plane and its sensitivity curve in *z* direction are shown in Fig.2(c) and 1(d), respectively. Compared to the sensitivity distribution without adjustment (shown in Fig.2 (a) and 2(b)), the sensitivity difference in Fig. 2(c) and 2(d) between deep and superficial layers has been reduced, which indicates that the sensitivity in deep layers has obtained some compensation and the depth resolution for the reconstructed image would be improved by this method.

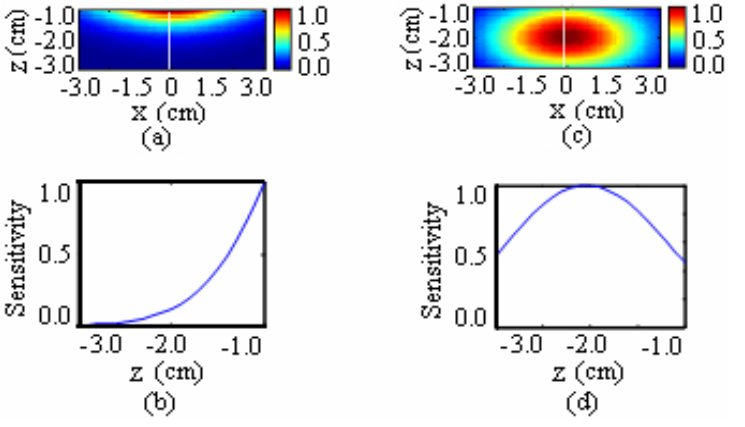


Fig. 2. Graphs of the overall sensitivity within x-z slice and the sensitivity curves with x=0 in the x-z slices. Of which (a) and (b) are the original sensitivity distributions and (c) and (d) are the sensitivity distributions using LMA technique.

4 Evaluation Criteria of Reconstructed Images

We use two criteria in this work to evaluate the reconstructed image quality: contrast to noise ratio (CNR) and position errors (PE). CNR was calculated using the following equation [16]:

$$\mathrm{CNR} = \frac{\mu_{ROI} - \mu_{ROB}}{\left[\omega_{ROI}\sigma_{ROI}^2 + \omega_{ROB}\sigma_{ROB}^2\right]^{1/2}} \tag{5}$$

where ω_{ROI} is the division of the area of the region of interest by the area of the whole image, and ω_{ROB} is defined as ω_{ROB}=1-ω_{ROI}, μ_{ROI} and μ_{ROB} are the mean value of the objects and background regions in reconstructed images. σ_{ROI} and σ_{ROB} are the corresponding standard deviations.

CNR indicates whether an object could be clearly detected in reconstructed images. PE is the distance between centers of the real object and the detected object. Greater CNR values and smaller PE values are considered as higher images quality.

5 Simulations and Results

To investigate practical applicability of the LMA method, we reconstruct images with objects located at depths of -1.0 cm, -1.3 cm, -1.6 cm, -1.9 cm, -2.1 cm, -2.4cm, -2.7 cm and -3.0 cm, as shown in the first column of Fig. 3. The reconstructed images without

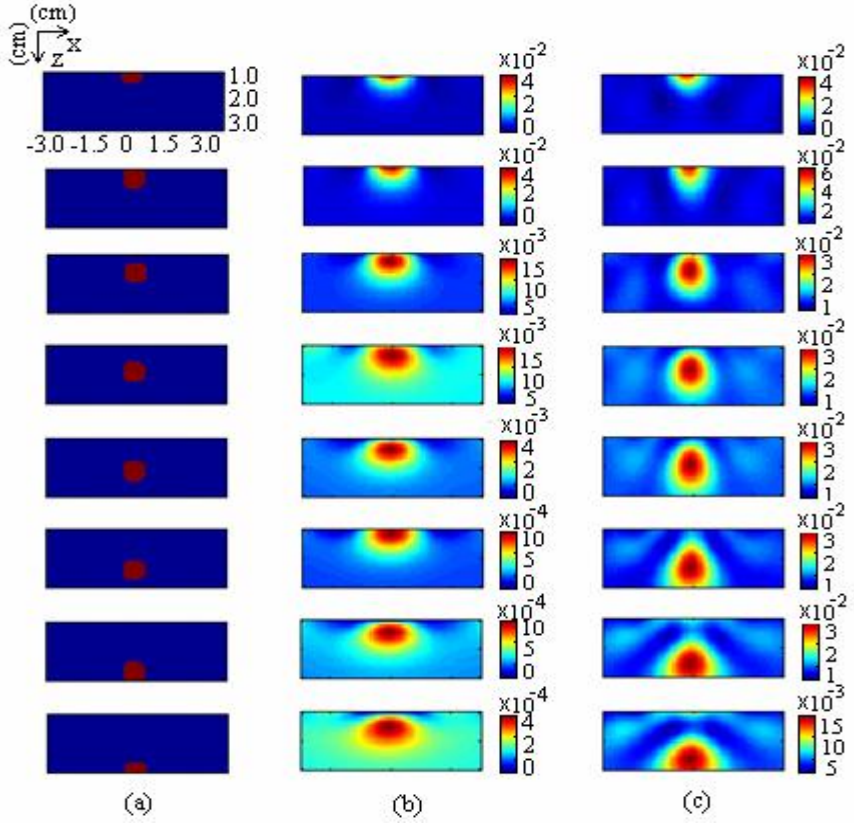


Fig. 3. (a) Original images, (b) reconstructed images without adjustment, and (c) with the LMA method. The objects are located at depths of -1.0 cm, -1.3 cm, -1.6 cm, -1.9 cm, -2.1 cm, -2.4 cm, -2.7 cm and -3.0 cm (from top to bottom).

Table 1. The CNR and PE values for the reconstructed images shown in Fig.3

Depth	Unadjusted		LMA	
(cm)	CNR	PE(mm)	CNR	PE(mm)
-1.0	16.21	0.64	17.51	0.46
-1.3	10.95	1.75	13.54	0.95
-1.6	8.31	2.26	11.17	0.98
-1.9	7.34	2.71	10.39	0.25
-2.1	3.92	4.91	9.46	0.73
-2.4	2.91	5.62	11.47	0.12
-2.7	0.39	11.81	8.02	0.87
-3.0	0.28	11.25	7.63	2.12

and with the LMA adjustment are shown in the second and third column of Fig.3, respectively. The CNR and PE values for the reconstructed images are listed in Table 1.

The reconstructed images without any adjustment in the forward matrix are clearly biased in favor of the superficial layers. This can be observed from the second column of Fig. 3. As a result, high CNR values and small PE values can be seen in Table 1 for depths from -1.0 cm to -1.9 cm. At the same time much smaller CNR values and much larger PE values can be found for depths from -2.1 cm to -3.0 cm. This indicates that DOT without any adjustment has poor performance in the depth layers.

In contrast, the images reconstructed with the LMA method can achieve a satisfactory image quality for both superficial and deep objects. Correspondingly, the larger CNR values and smaller PE values are obtained, as shown in Table 1. In particular, PE less than 3.0 mm can be obtained for all the images reconstructed with the LMA when the object was located from -1.0 cm to -3.0 cm. This indicates a significant improvement in depth resolution by the LMA method compared to the original reconstructed images.

6 Conclusion

In this paper, we demonstrated that the depth resolution of DOT imaging can be significantly improved by applying the intelligent approach in the sensitivity matrix. Simulations implemented in a semi-infinite model indicate that the LMA method can effectively reduce the sensitivity contrast between the deep and superficial tissues. The larger CNR values and smaller PE values of the reconstructed images demonstrate that the LMA method is very promising in improving the image quality. It should also be noted that a second advantage for the LMA method, it is more convenient to apply due to no additional empirical parameters in it.

Acknowledgements. This work was supported by the Natural Science Foundation of China, Grant Nos. 60675011 and 30425004, the National Key Basic Research and Development Program (973), Grant No. 2003CB716100, and the National High-tech R&D Program (863 Program) No.2006AA01Z132.

References

1. Pogue, B.W., Testorf, M., Mcbride, T., et al.: Instrumentation and Design of a Frequency Domain Diffuse Optical Tomography Imager for Breast Cancer Detection. Opt. Exp. 13, 391–403 (1997)
2. Hebden, J.C., Gibson, A., Yusof, R.M., Everdell, N., et al.: Three-dimensional Optical Tomography of the Premature Infant Brain. Phys. Med. Biol. 47, 4155–4166 (2002)
3. Bluestone, A.Y., Abdoulaev, G., et al.: Three Dimensional Optical Tomography of Hemodynamics in the Human Head. Opt. Exp. 9, 272–286 (2001)
4. Peters, V.G., Wyman, D.R., et al.: Optical Properties of Normal and Diseased Human Breast Tissues in the Visible and Near Infrared. Phys. Med. Biol. 35, 1317–1334 (1990)
5. Douiri, A., Schweiger, R.M.J., Arridge, S.: Local Diffusion Regularization Method for Optical Tomography Reconstruction by Using Robust Statistics. Opt. Lett. 30, 2439–3441 (2005)
6. Joseph, D.K., Huppert, T.J., et al.: Diffuse Optical Tomography System to Image Brain Activation with Improved Spatial Resolution and Validation with Functional Magnetic Resonance Imaging. Appl. Opt. 45, 8142–8151 (2006)
7. Niu, H.J., Guo, P., Ji, L., Jiang, T.: Improving Diffuse Optical Tomography Imaging with Adaptive Regularization Method. In: Proc. SPIE, vol. 6789 K, pp. 1–7 (2007)
8. Pogue, B.W., McBride, T.O., et al.: Spatially Variant Regularization Improves Diffuse Optical Tomography. App. Opt. 38, 2950–2961 (1999)
9. Zhao, Q., Ji, L., Jiang, T.: Improving Depth Resolution of Diffuse Optical Tomography with a Layer-based Sigmoid Adjustment Method. Opt. Exp. 15, 4018–4029 (2007)
10. Paulsen, K.D., Jiang, H.: Spatially Varying Optical Property Reconstruction Using a Finite Element Diffusion Equation Approximation. Med. Phys. 22, 691–701 (1995)
11. Ishimaru, A., Schweiger, M., Arridge, S.R.: The Finite-element Method for the Propagation of Light in Scattering Media: Frequency Domain Case. Med. Phys. 24, 895–902 (1997)
12. Delpy, D.T., Cope, M., et al.: Estimation of Optical Pathlength Through Tissue From Direct Time of Flight Measurement. Phys. Med. Biol. 33, 1433–1442 (1988)
13. Zhao, Q., Ji, L., Jiang, T.: Improving Performance of Reflectance Diffuse Optical Imaging Using a Multiple Centered Mode. J. Biomed. Opt. 11, 064019, 1–8 (2006)
14. Arridge, S.R.: Optical Tomography in Medical Imaging. Inverse Probl. Eng. 15, R41-R93 (1999)
15. Boas, D.A., Dale, A.M.: Simulation Study of Magnetic Resonance Imaging-guided Cortically Constrained Diffuse Optical Tomography of Human Brain Function. Appl. Opt. 44, 1957–1968 (2005)
16. Song, X., Pogue, B.W., et al.: Automated Region Detection Based on the Contrast-to-noise Ratio in Near-infrared Tomography. Appl. Opt. 43, 1053–1062 (2004)

Research on Optimum Position for Straight Lines Model*

Li-Juan Qin, Yu-Lan Hu, Ying-Zi Wei, Hong Wang, and Yue Zhou

School of Information Science and Engineering, Shenyang Ligong University, Shenyang, Liaoning Province, China
qinlijuan1978@gmail.com

Abstract. Quantization errors are the primary source that affects the accuracy of pose estimation. For the model at different placement positions, quantization errors have different effects on the results of pose estimation. The analysis of optimum displacement for the model can increase the accuracy of pose estimation. In this paper, mathematical model of the propagation of quantization errors from a two-dimensional image plane to the 3D model is set up. What's more, optimization function for the analysis of optimum placement position of model is set up. For given model, we obtain optimum placement position of model with respect to camera. At last, the simulation results show that it has better pose estimation accuracy at optimum place than at other places.

Keywords: Pose estimation; Quantization errors; Model; Placement position.

1 Introduction

Model-based monocular pose estimation is an important problem in computer vision. It has wide varieties of applications, such as robot self-positioning, robot navigation, robot obstacle avoidance, object recognition, object tracking, hand-eye calibration and camera calibration [1]. At present, the pose estimation methods include line-based pose estimation methods and point-based pose estimation methods. The accuracy of pose estimation results depends critically on the accuracy with which the location of the image features can be extracted and measured. In practical application, line feature is easier to extract and the accuracy of line detection is higher than point because the images we obtain are often unclear and occluded [2]. Thus pose estimation results of line correspondences are usually more robust than that of point correspondences.

Quantization errors are the primary sources that affect the precision of pose estimation and they are inherent and unavoidable. For the quantization errors in the model-based pose estimation system, previous research has introduced some results on spatial quantization errors. Kamgar-Parsi [3] developed the mathematical tools for computing the average error due to quantization. Blostein [4] analyzed the effect of image plane quantization on the three-dimensional (3-D) point position error obtained by triangulation from two quantized image planes in a stereo setup. Ho [5] expressed the digitizing error for various geometric features in terms of the dimensionless

* Supported by National High Technology Research and Development Program of China under grant No. 2003AA411340.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 521–528, 2008.
© Springer-Verlag Berlin Heidelberg 2008

perimeter of the object, and Griffin [6] discussed an approach to integrate the errors inherent in the visual inspection. The spatial quantization error of a point in one dimension has been modeled as a uniform distribution in previous work. Using a similar representation, paper [7] analyzes the error in the measured dimension of a line in two dimensions. The mean, the variance and the range of the error are derived. Quantization errors are not constant in the pose estimation process. It depends on the computed value itself. Here, we shall discuss the way the error propagates for line-based pose estimation method. This paper discusses the effect of the quantization errors on the precision of dimensional measurements of lines. At present, there is little research on how these quantization errors propagate from a two-dimensional (2-D) image plane to the three-dimensional (3-D) model according to the literatures we have referred. Therefore, it is important that these errors can be analyzed. Here, different from predecessors, we analyze how the quantization errors affect the accuracy of pose estimation. At the same time, we set up the mathematical model. The analysis of the effect of quantization errors of image lines on pose estimation results from line correspondences can estimate the error range of pose estimation system and direct the placement of location model.

The placement of location model can be arranged in advance. In model-based monocular pose estimation system, for a given model, quantization errors have different effects on the model at different placement positions. Thus, for different placement places of the model, the accuracy of pose estimation is different. Thus, the analysis of the optimum placement for the model is an important aspect to execute the vision task at high precision. At present, there is little research on the optimum placement of the model with respect to the camera according to the literatures we have referred. On the basis of analysis of the propagation of the quantization errors, we do the research on the optimum placement of the model with respect to camera. The purpose of the research is to increase the accuracy of pose estimation.

The remainder of this paper is organized as follows: In section 2, pose estimation algorithm from line correspondences is presented. Section 3 introduces the propagation of quantization errors. In section 4, the method to determine the optimum placement for the model is presented. In section 5, experimental results are presented. Section 6 is the conclusion.

2 Problem Formulation

At present, most research is focus on the pose estimation from three line correspondences, which is called “P3L”. This problem can be described as follows: we know the 3D line equations in the object frame and we know the corresponding line equations in the image plane. The problem is to determine the rigid transformation (rotation and translation) between the object frame and the camera frame (see Fig.1)[8].

We consider a pin-hole camera model and we assume that the parameters of the projection (the intrinsic camera parameters) are known. As shown in Fig.1, we assume the vectors of model line L_i （ $i=1,2,3$ ） is $V_i(A_i,B_i,C_i)$. We consider an image line characterized by a vector v_i $(-b_i,a_i,0)$ and a point t_i of coordinates of

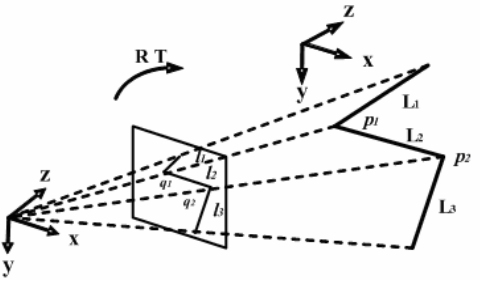


Fig. 1. Perspective projection of three lines

(x_i, y_i, f) .The perspective projection model constraints line L_i , image line l_i and the origin of the camera frame to lie in the same plane. This plane is called explanation plane. The vector N_i normal to this plane can be computed as the cross product of ot_i and vector v_i , we have:

$$ot_i = \begin{pmatrix} x_i \\ y_i \\ f \end{pmatrix}, \quad v_i = \begin{pmatrix} -b_i \\ a_i \\ 0 \end{pmatrix}$$

Then:

$$N_i = ot_i \times v_i = \begin{pmatrix} a_i f \\ b_i f \\ c_i \end{pmatrix} \tag{1}$$

We assume the vectors of space line L_i in the object frame are $n_i = (A_{wi}, B_{wi}, C_{wi})$ .The rotation matrix between the camera frame and the object frame is R . The space line in the camera frame after the rotation is V_i , then we obtain:

$$V_i = Rn_i \tag{2}$$

From the relationship that the normal of the explanation plane is perpendicular to the space line, we can get the equation about rotation matrix R as follows [9]:

$$N_i \cdot Rn_i = 0 \tag{3}$$

Thus, we can get the rotation parameters of the rotation matrix R from three image lines. To solve the rotation matrix, there are not two lines which are parallel, and lines do not pass the optical centre.

3 Propagation of Quantization Errors

Because of quantization errors, errors of image lines will lead to errors of rotation angles α , β , γ . In the following, we will do the research on the relationship between the quantization errors and the errors of α , β , γ .We assume the error of the norm

vector $N_i = (N_{i1}, N_{i2}, N_{i3})$ of the explanation plane is dN_i because of quantization errors. The error of the pose estimation result is $dR = (d\alpha, d\beta, d\gamma)$ because of the effect of quantization errors.

From the pose estimation method from three line correspondences, we obtain:

$$N_i \cdot Rn_i = 0$$

$$R = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \tag{4}$$

$$n_i = \begin{pmatrix} A_{wi} \\ B_{wi} \\ C_{wi} \end{pmatrix} \quad N_i = \begin{pmatrix} N_{i1} \\ N_{i2} \\ N_{i3} \end{pmatrix} \tag{5}$$

Substituting (4) and (5) into (3), We can obtain equations between α, β, γ and $(N_{i1}, N_{i2}, N_{i3})(i = 1, 2, 3)$:

$$\begin{cases} F_1(N_{11}, N_{12}, N_{13}, \alpha, \beta, \gamma) = 0 \\ F_2(N_{21}, N_{22}, N_{23}, \alpha, \beta, \gamma) = 0 \\ F_3(N_{31}, N_{32}, N_{33}, \alpha, \beta, \gamma) = 0 \end{cases} \tag{6}$$

We assume $N = [N_{11}, N_{12}, N_{13}, N_{21}, N_{22}, N_{23}, N_{31}, N_{32}, N_{33}]$, $R = [\alpha, \beta, \gamma]$, $F = [F_1, F_2, F_3]$, we get the partial differential equation for (6):

$$\frac{\partial F}{\partial N} dN + \frac{\partial F}{\partial R} dR = 0 \tag{7}$$

Thus, the relationship between the error dN of N and error dR of R is:

$$dR = -\left(\frac{\partial F}{\partial R}\right)^{-1} \left(\frac{\partial F}{\partial N}\right) dN \tag{8}$$

It can be arranged into:

$$\begin{cases} f_{11}d\alpha + f_{12}d\beta + f_{13}d\gamma + f_{14}dN_{11} + f_{15}dN_{12} + f_{16}dN_{13} = 0 \\ f_{21}d\alpha + f_{22}d\beta + f_{23}d\gamma + f_{24}dN_{21} + f_{25}dN_{22} + f_{26}dN_{23} = 0 \\ f_{21}d\alpha + f_{22}d\beta + f_{23}d\gamma + f_{24}dN_{21} + f_{25}dN_{22} + f_{26}dN_{23} = 0 \end{cases} \tag{9}$$

From (9), we have:

$$d\gamma = \frac{s_2(f_{12}f_{21} - f_{22}f_{11}) - s_1(f_{22}f_{31} - f_{32}f_{21})}{(f_{13}f_{21} - f_{23}f_{11})(f_{22}f_{31} - f_{32}f_{21}) - (f_{23}f_{31} - f_{33}f_{21})(f_{12}f_{21} - f_{22}f_{11})} \tag{10}$$

$$d\beta = \frac{(f_{13}f_{21} - f_{23}f_{11})d\gamma + s_1}{-f_{12}f_{21} + f_{22}f_{11}} \tag{11}$$

$$d\alpha = \frac{-(f_{12}d\beta + f_{13}d\gamma + f_{14}dN_{11} + f_{15}dN_{12} + f_{16}dN_{13})}{f_{11}} \tag{12}$$

where, Equation（10）（11）（12）analysis the way the quantization errors propagate and obtain the relationship between the errors of α , β , γ and the errors of N_i in the form of closed-form solutions. We assume the errors of the norm vectors of line features have been modeled as a uniform distribution.

We set up the optimization function for the optimum placement of the model with respect to the camera as follows:

$$F = |d\alpha| + |d\beta| + |d\gamma| \tag{13}$$

When the sum of the errors for the pose estimation be minimum, we can determine the optimum place for the model with respect to the camera. In practical application, we can use it to determine the optimum placement for given model with respect to camera. Thus it can increase the pose estimation accuracy.

4 Determination of Optimum Placement Position

For given model (see Fig. 2), the quantization errors have different effects on the model at different placement positions. (Quantization errors have been modeled as a uniform distribution)

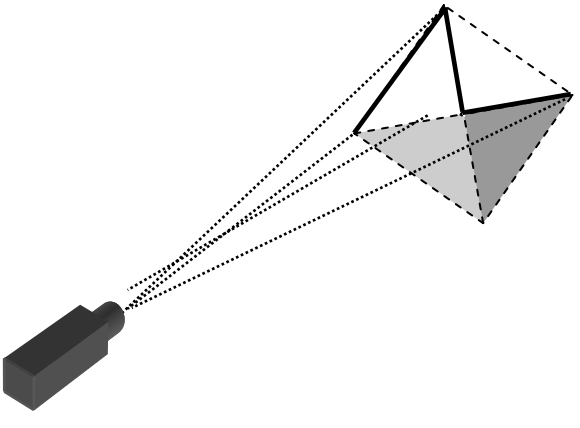


Fig. 2. Original placement position of model

For different placement positions of the model, we input the same quantization errors and analyze the value of function F. The placement position of the model corresponding to the minimum value of function F is the optimum placement position.

Here, we use global strategy to find optimum placement position. For given model (see Fig. 2), we make the model rotate around z axis at the range of $(-180°,180°)$, rotate around x axis at the range of $(-30°,30°)$, and rotate around y axis at the range of $(-30°,30°)$ at the step of $1°$. We set up different testing places for the model and obtain the value of function F at these testing places.

For this given model, the simulation results show that the sum of errors is small when the model rotate around z axis of $45°$ (see Fig. 3). The final results show the accuracy of pose estimation is higher at this place than at other places.

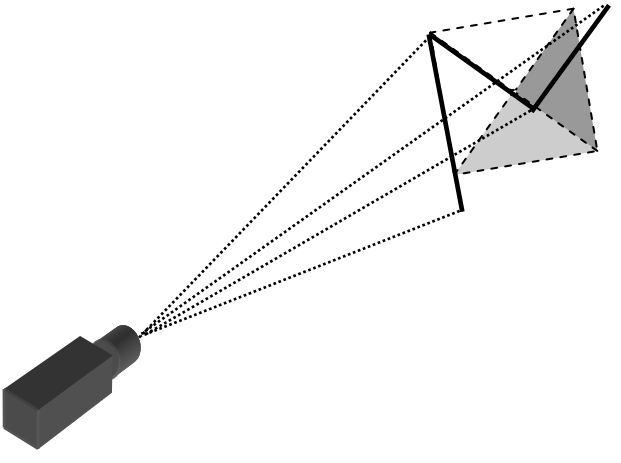


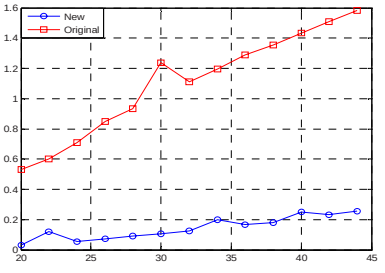
Fig. 3. Optimum placement position of model

5 Experimental Results

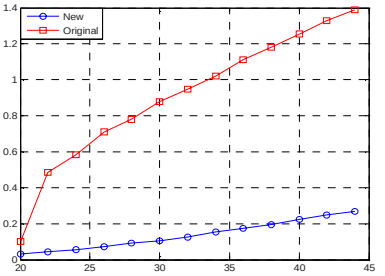
In our simulation experiments, we choose the concrete parameters of the camera are $\begin{pmatrix} 787.887 & 0 & 0 \\ 0 & 787.887 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The size of image is 512×512. The view angle of camera is about 36×36 degrees.

We add 0.01pixel random noise to the norm vector of the image lines. We do some statistics about RMS error. At last, the rotation error is equal to three times of the RMS orientation error. In Fig.4, the horizontal coordinates are the ratio of the distance between the optical center and the model to the model size. The vertical coordinates of Fig.4（a-c）are degrees. The model size is 15mm.

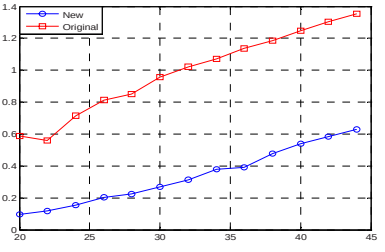
We do the contrastive experiments with the initial placement position. The results of pose estimation are as follows:



(a) Orientation error in x axis



(b) Orientation error in y axis



(c) Orientation error in z axis

Fig. 4. Comparison result of errors

From simulation results, it shows that the accuracy of pose estimation have been improved greatly when the model rotate around z axis of $45°$.

6 Conclusion

For pose estimation from line correspondences, we analyze the characteristics for propagation of quantization errors and set up the mathematical model for the effect of quantization errors on the accuracy of pose estimation. For a given model, quantization errors have different effects on location accuracy of model at different placement positions. The analysis of optimum placement of location model with respect to camera is an important aspect to execute vision task at high precision. This paper sets up optimization function for optimum placement position of location model. Furthermore, we use global strategy to find optimum placement position by finding the minimum value of optimization function. The simulation results show the accuracy of pose estimation has been increased greatly at optimum place. The method can be used to determine the optimum placement position for an arbitrary model. Thus, this analysis can increase the accuracy of pose estimation system.

References

1. Lee, Y.R.: Pose Estimation of Linear Cameras Using Linear Feature. The Ohio State University, Dissertation (2002)
2. Christy, S., Horaud, R.: Iterative Pose Computation from Line Correspondences. Computer Vision and Image Understanding 73(1), 137–144 (1999)
3. Kamgar-Pars, B.: Evaluation of Quantization Error in Computer Vision. IEEE Trans. Pattern Anal. Machine Intell. 11(9), 929–940 (1989)
4. Blostein, S.D., Huang, T.S.: Error Analysis in Stereo Determination of 3-D Point Positions. IEEE Transactions on Pattern Analysis and Machine Intelligence 9(6), 752–765 (1987)
5. Ho, C.: Precision of Digital Vision Systems. In: Workshop Industrial Applications of Machine Vision Conference, pp. 153–159 (1982)
6. Griffin, P.M., Villalobos, J.R.: Process Capability of Automated Visual Inspection Systems. IEEE Trans. Syst., Man, Cybern. 22(3), 441–448 (1992)
7. Christopher, C.Y., Michael, M.M.: Error Analysis and Planning Accuracy for Dimensional Measurement in Active Vision Inspection. IEEE Transactions on Robotics and Automation 14(3), 476–487 (1998)
8. Dhome, M., et al.: Determination of the Attitude of 3-D Objects from a Single Perspective View. IEEE Trans. Pattern Anal. Machine Intell. 11(12), 1266–1278 (1989)
9. Liu, Y., Huang, T.S., Faugeras, O.D.: Determination of Camera Location from 2-D to 3-D Line and Point Correspondences. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(1), 28–37 (1990)

Choosing Business Collaborators Using Computing Intelligence Methods

Yu Zhang, Sheng-Bo Guo, Jun Hu, and Ann Hodgkinson

The University of Wollongong,
Wollongong, NSW 2522, Australia
{yz917,sbg774,jh928,annh}@uow.edu.au

Abstract. Inter-firm collaboration has become a common feature of the developing international economy. Firms as well as the nations have more relationships with each other. Even relatively closed economies or industries are becoming more open, Australia and China are examples of this case. The benefits generated from collaboration and the motivations to form collaboration are investigated by some researchers. However, the widely studied relationships between collaboration and profits are based on tangible assets and turnovers whereas most intangible assets and benefits are neglected during the economic analysis. In the present paper, two methods, naive Bayes and neural network, from computing intelligence are used to study the benefits acquired from collaboration. These two methods are used to learn the relationship and make prediction for a specified collaboration. The proposed method has been applied to a practical case of WEMOSOFT, an independent development department under MOBOT. The predication accuracies are 87.18% and 92.31%, for neural network and naive Bayes, respectively. Experimental result demonstrates that the proposed method is an effective and efficient way to prediction the benefit of collaboration and choose the appropriate collaborator.

Keywords: Business Collaborators; Computing Intelligence; MOBOT.

1 Introduction

In recent years, the economic climate has fostered industrial cooperation more strongly than in the past. Cooperation has often proven superior to outright competition [1]. "Unless you've teamed up with someone, you become more vulnerable as barriers fall. The best way to eliminate your potential enemies is to form an alliance with them." [2].

Many researchers in different fields have decoded the principles of collaborating and cooperating. They have attempted to solve the problems and conflicts in collaborations and increase the benefits to all partners who participate in such collaborations. Countries, governments, firms, educators and even individuals benefited from their research.

Coase recognized the important role of Transaction Cost as well as the role of firms and argued that transaction costs were the reason why firms exist [3]. Arrow also contributed to the appreciation of the role of Transaction costs [4]. The main approach in Transaction Cost Economics is proposed in [5]. In addition, Williamson further categorised inter-firm transactions into competition (market transaction), governance (internal transaction), planning (contract), and promise (collaboration) [6]. Contractor et

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 529–535, 2008.
© Springer-Verlag Berlin Heidelberg 2008

al. believed that cooperation and competition provide alternative or simultaneous paths to success [1]. Much of the recent research is based on network cooperation, new technologies and multinational enterprises. For example: Hagedoorn's research on technology partnership [7] [8] [9]; Gilroy's work on the networking benefits for multinational enterprizes [10]; Roos's paper on the cooperating strategies for global business [11]; Kay's research on innovation and trust [12]; Chen and Shih's research on high-tech development [13]; Zhang and Dodgson's research on telecommunication cooperation in Asia [14].

These researchers recognized collaboration in different fields and industries. However, with globalization, the role and types of collaboration are changing rapidly. The risks facing different firms for collaboration have also changed due to emerging new technology and markets.

This paper discusses why firms collaborate; how they collaborate; and what are the main risks facing their collaborations. The motives for collaboration will be used to develop the questionnaire to collect data from firms and build the model. Different types of collaborations will help to identify the rapid changes in global market and collaboration. To increase the inter-firm collaboration, it is necessary to reduce or eliminate the risks associated with collaboration. Finally, the results of the case study will help in finding a new solution for inter-firm collaborations.

To understand better of collaboration, increased endeavor has been put into analysis and prediction of the result from collaboration. Conventionally, the analysis is based on manual, and thus exhaustive. In this paper, we propose to analyze the results of collaboration by using intelligent computing. Two models are proposed to develop the experience of collaboration, and are then used for predicting the result for future collaboration.

The remainder of this paper is organized as follows. Section 2 presents the motives, types and risks of collaboration. Computing intelligence methods for analyzing the results of collaboration are formulated in Section 3. Section 4 present the case study of WEMOSOFT. Section 5 is the conclusion, followed by future work.

2 The Motives, Types and Risks of Collaboration

The incentives for firms to collaborate may be external or internal. They all target the impact on final net profits or utilities, but with different business and political environments. Collaborations show great variety in their motivation. The external reasons or pressure that made firms collaborate may include: rapid economic and technological change; declining productivity growth and increasing competitive pressures; global interdependence; blurring of boundaries between business, overcoming government-mandated trade or investment barriers; labor; lack of financial resources; facilitating initial international expansion of inexperienced firms; and dissatisfaction with the judicial process for solving problems [1] [15] [16]. On the other hand, there are some benefits generated from collaborating, which may push firms that are chasing profits to collaborate: to access producers of information, new markets, benefits to the consumer;

lower coordination costs throughout the industry value chain; lower physical distribution costs; redistribution and potential reduction in total profits; reduction of innovation lead time; technological complementary; influencing market structure; rationalization of production; monitoring technological opportunities; specific national circumstances; basic R&D and vertical quasi-integration advantages of linking the complementary contributions of the partners in a ”value chain” [1] [16].

Pfeffer et al [17] and later Contractor et al [1] categorized the major types of cooperative as: Technical training and start-up assistance agreements; Production, assembly, and buy-back agreements; Patent licensing; Franchising; Know-how licensing; Management and marketing service agreement; Non equity cooperative agreements in Exploration, Research, partnership, Development, and co-production; and Equity joint venture.

Risk levels increase as the depth of cooperation increases: historical and ideological barriers; power disparities; societal-level dynamics creating obstacles to collaboration; differing perceptions of risk; technical complexity; political and institutional cultures; the role of management; the benefit distribution; and so on [15].

With the dynamic changes in global markets, it is hard to identify whether a company is safe when contributing to a new collaboration or not. Even the big companies can not category ”good” and ”bad” cooperators. The ”bad” cooperators here are interpreted as those cooperators that occupied some of resources but didn’t contributed any tangible and intangible benefit. The learning process is a huge work beyond one person’s ability. It needs thousands of real cases to clarify the blurred boundary from ”good” to ”bad”. One possible way to analyze these collaborations are to use the methods from computing intelligence, however, these are rarely reported when analyzing the results of collaboration. Motivated by the advanced learning and member ability, we propose to employ two methods, namely, naive Bayes and neural network for analyzing business collaboration.

3 Methods

To investigate the benefits from collaboration, we propose to use two models, namely, naive Bayes and neural networks [18]. Naive Bayes classifier is a simple probabilistic classifier based on applying Bayes’ theorem with naive independence assumptions. It is designed for use in supervised induction task, in which the performance goal is to accurately predict the class of test instances and in which the training instances include class information [19].

Aside from naive Bayes classifier, neural network (NN) is also employed for prediction because NN is used to simulate the network formed by neurons in brain. Specifically, We use multi layer perceptron (MLP), and train the MLP using backpropagation algorithm. The structure of the MLP is shown in Fig. 1.

To evaluate the performance of predication, the Leave-One-Out Cross-Validation (LOOCV) is conducted, and the predication accuracy is report based on LOOCV. The experiment is conducted using the Waikato Environment for Knowledge Analysis (WEKA) [20].

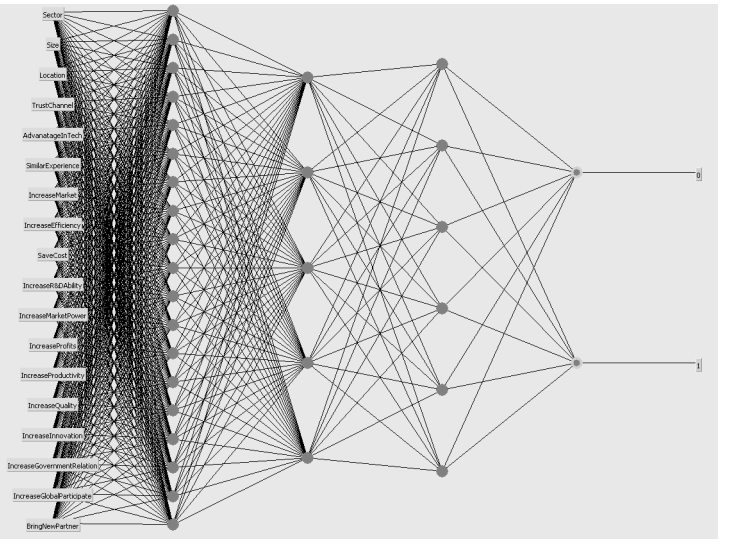


Fig. 1. The structure of the MLP for prediction of benefit from collaboration

4 Case Study

WEMOSOFT is an independent development group under Beijing MOBOT Software Technology Co., Ltd., which was established in 2005 and have successfully implemented more than ten projects in different fields. All the successful projects are based on good collaboration and cooperation. WEMOSOFT established its business network with more than one hundred telecommunication and software companies during 2005 to 2007. Some of them contributed directly to the successful projects and tangible profits, some contributed to its intangible benefits, but there are some firms didn't and seemingly won't contribute to any business operation. As a result, we separated all firms in the business network into three groups: tangible contributors, intangible contributors and none contributors. Due to the data analysis, there are about 1/10 tangible contributors and 1/3 intangible contributors in all business networks, which implies there are more than 2/3 none contributors but occupied most of the business resources and efforts. The purpose of this paper is how to use computer intelligence to reduce these costs and efforts in business operation.

4.1 Data Set Collection

The researched data was collected from WEMOSOFT, a Chinese company located in Beijing. WEMOSOFT was involved in business collaboration with its 39 collaborators from mid 2000 to the end of 2006. The researcher collected all the major features of its business collaboration via a designed questionnaire. All the data are analyzed based on the business collaboration categories included in part two and research conducted by the researchers.

There are 18 major determinants that will influence the success of business collaboration. They are: business sector, which described whether the collaborator is in the same or similar business sector of the researched company.

The determinants are defined as follow. Business size described the size similarity of WEMOSOFT and its collaborator. Business location is the geographic distance of the company, which also related to transaction costs for this collaboration. Business trust channel describes the level of trust before the collaboration. It depends on the business network and trust of WEMOSOFT, how it is known to the collaborator, how many layers of business relationship between it and its collaborator. This feature may be very complex, but all the data are virtualized into quantization for the convenience of research. The advantage in technology described the technical level of the collaborator, which is supposed to have a positive influence on the success of collaboration. Similar experience describes whether the collaborator had similar collaborating experience with another company before this collaboration. A positive relationship between a successful collaboration and increased market, increased efficiency, cost saving, increased R&D (research and development) ability, increased market power, increased profits, increased productivity, increased quality, increased innovation, improved government relationship, increased global participation, and bring new partners are expected.

The collected attributes are first normalized into the $[0, 1]$ interval according to the related influence, where $'0'$ indicates that this determinant has no contribution to the final result, $'1'$ indicates that this determinant has strong contribution to the final result. For class label, it is set to $'0'$ if the collaboration failed, otherwise, it is set to $'1'$. By doing so, we develop the data sets X with 39 rows and 18 columns, and Y with 39 rows and 1 column. For X, its ith row represents the ith sample with its class label represented in the ith row in Y.

4.2 Experimental Results

By using LOOCV, the prediction accuracy by using the naive Bayes classifier and MLP is 92.31% and 87.18%. The confusion matrix for naive Bayes and MLP are shown in Table 1 and Table 2, respectively. The confusion matrix indicated that both models achieved the same performance for predicting successful collaboration, whereas the naive Bayes classifier is superior to the MLP when classifing the failed collaborations.

Table 1. The confusion matrix of naive Bayes for predicting the benefit of collaboration

Collaboration Result	Success	Fail
Success	21	2
Fail	1	15
Accuracy	95.45%	88.24%

In summary, the proposed method achieved satisfactory performance for predicting the result of collaboration. As a consequence, this approach proved to be effective for analyzing the benefit of collaboration in the case study.

Table 2. The confusion matrix of MLP for predicting the benefit of collaboration

Collaboration Result	Success	Fail
Success	21	2
Fail	4	12
Accuracy	84%	85.71%

5 Conclusions

This paper proposes an alterative way to estimate the results of collaboration and cooperation based on computing intelligence, which plays an important role in the development of firms. Collaboration and cooperation is a way to generate more profits but also a vital strategy for most firms experiencing fast growth associated with globalization and international competition. Different firms in different circumstances may have varied motives in forming collaborations. With the development of the economic activities and variety of new firms as well as new markets, the type of collaboration is increasing and changing as well. Some examples are given to different types of collaboration. Firms as well as the institutions are searching for methods to reduce the risks from collaboration. It is more important for small and medium sized firms to plan and adopt collaboration strategies to survived fierce competition in the future. The proposed approach can be used to assist the analyze of making decision for collaboration.

6 Future Work

Future work will investigates the influence of each attribute on result of collaboration by using computing intelligence. By doing so, the key factors that lead to the success of collaboration can be identified. These factors can be use to improve the corresponding attributes of a firm for the success of collaboration with another firm. Moreover, advanced computing intelligence methods will also be developed for accurate prediction.

References

1. Contractor, F.J., Lorange, P.: Cooperative Strategies in International Business. Lexington Books, Canada
2. International Herald Tribune: Sprint Deal with Europe Sets Stage for Phone War (1994)
3. Coase, R.H.: The Nature of the Firm. Economics 4, 386–405 (1937)
4. Arrow, K.J.: The Organization of Economic Activity: Issue Perinent to the Choice of Market Versus Nonmarket Allocation. The analysis and evaluation of public Expenditure: The PPB System 1, 59–73 (1969)
5. Williamson, O.E.: The Economic Institutions of Capitalism. The Free Press, New York (1985)
6. Williamson, O.E.: Handbook of Industrial Organization. In: Transaction Cost Economics, pp. 136–182. Elsevier Science, New York (1989)
7. Hagedoorn, J.: Understanding the Rationale of Strategic Technology Partnering: Interorganizational Modes of Cooperation and Sectoral Differences. Strategic Management Journal 14, 371–385 (1993)

8. Hagedoorn, J.: Research Notes and Communications a Note on International Market. Strategic Mangement Journal 16(3), 241 (1995)
9. Hagedoorn, J., Hesen, G.: Understanding the Cross-Level Embeddedness of Interfirm Partnership Formation. Academy of management review 31(3), 670–680 (2006)
10. Gilroy, B.M.: Networking in Multinational Enterprises - The Importance of Stratetic Aliances, University of South Carolina, Columbia, South Carolina
11. Roos, J.: Cooperative Strategies. Prentice Hall, UK (1994)
12. Kay, N.M.: The boundaries of the firm - Critiques, Strategies and Policies. Macmillan Press Ltd., Basingstoke (1999)
13. Chen, C.H., Shih, H.T.: High-Tech Indistries in China. Edward Elgar, UK (2005)
14. Zhang, M.Y., Dodgson, M.: High-Tech Entrepreneurship in Aisa - Innovation, industry and institutional dynammics in mobile payments. Edward Elgar, UK (2007)
15. Gray, B.: Collaborating - finding common ground fro multiparty problems. Jossey-Bass Publisher, San Franciso (1998)
16. Freeman, C., Soete, L.: New Explorations in the Economics of Technical Change. Printer Publisher, London, New York (1990)
17. Pfeffer, J., Nowak, P.: Joint Ventures and Interorganizational Interdependence. Administrative Science Quarterly 21, 398–418 (1976)
18. Duda, R.O., Hart, P.E., Stork, D.G.: Patttern Classification, 2nd edn. John Wiley & Sons, Inc., Chichester (2001)
19. John, G.H., Langley, P.: Estimating Continuous Distributions in Bayesian Classifiers. In: Proceedings of the Eleventh Conference on Uncertainty in Artificial Inteilligence, San Mateo, pp. 338–345. Morgan Kaufmann Publisher, San Francisco (1995)
20. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques, 2nd edn. Morgan Kaufmann, San Francisco (2005)

Generation of Multiple Background Model by Estimated Camera Motion Using Edge Segments

Taeho Kim and Kang-Hyun Jo

Graduate School of Electrical Engineering, University of Ulsan,
San 29, Mugeo-Dong, Nam-Gu, Ulsan, 680 - 749, Korea
{thkim,jkh2008}@islab.ulsan.ac.kr

Abstract. We investigate new approach for segmentation of moving objects and generation of MBM (Multiple Background Model) from an image sequence by a mobile robot. For generating MBM from unstable camera, we have to know the camera motion. When we correlate two consecutive images to calculate the similarity, we use edge segments to reduce computational cost. Because the regions, neighbors of edge segments, have distinctive spatial features while some regions like blue sky, empty road, etc. have ambiguity. Based on the similarity result, we obtain best matched regions, their centroids and displacement vector between two centroids. The highest density of displacement vector histogram, named motion vector, indicates camera motion between consecutive frames. We generate MBM based on motion vector and MBM algorithm classifies each matched pixel to several clusters. The experimental results shows that proposed algorithm successfully detect moving objects with MBM when camera has 2-D translation.

Keywords: MBM(Multiple Background Models), moving camera, object detection, Edge segment features.

1 Introduction

Human supporting or high security system is one of the main tasks in HCI(Human Computer Interaction). Furthermore the necessity of those systems has been proportionally increased with science technology. The examples of those systems are the video surveillance system and automobile robot. For the video surveillance or the automobile robot, it is important to segment moving objects. The common feature of two examples is the view point of image sequence is unstable. Therefore we suggest an idea "How to generate background model while camera moves" in this paper. Automatically generated background model has good merits for moving robot. It supports moving robot to select landmarks and detect moving objects using background subtraction. Generation of robust background is a part of main tasks in image analysis. Currently, many methods of generating background are already developed. Several examples of those are A single Gaussian background model by Wren et al. [1], MOG(Mixture of Gaussian) by Stauffer et al. [3] and multiple background images by Xiao et al. [10]. There merits and demerits are explained in [12]. However their methods are dealt in static camera, it is different with our purpose.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 536–543, 2008.
© Springer-Verlag Berlin Heidelberg 2008

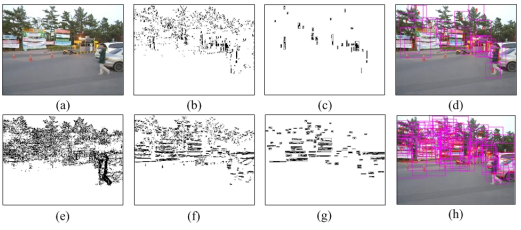


Fig. 1. Selected edge segment features and searching area

Moreover, moving object detection by moving camera is widely researched however their focuses are detecting moving objects itself. [6, 8, 11, 12]. In this paper, we propose a background reconstruction method which effectively constructs the background for dynamic scene and moving camera. For this reason we have to define the camera motion to three categories.

1. 2-Dimensional translation. (top, bottom, left, right)
2. Rotations based on each camera axis
3. Translation to the optical axis

When we record image by hand, the cases of 1∼3 are simultaneously occurred. Proposed algorithm in this paper describes how to calculate the camera displacement and generate multiple background model for the case 1. Furthermore we have been researching the algorithms to overcome the problems of calculating camera displacement for case 2 and 3.

2 Camera Motion

We easily detect moving object if we generate robust background. As we know in the recent research [1,2,3,4,5,9,10,12], many algorithms to generate background are already existed. Unfortunately, most algorithm are weaken when camera has unexpected effect like shaking, waving or moving for a long time. In this paper, we present pixel based multiple background algorithm that is calculating similarity between consecutive frames.

2.1 Feature Based Similarity

It is necessary to calculate camera motion if we generate background model by moving camera. However it needs high cost to calculate the camera motion using correlation between current and prior images. Therefore we choose not corner points but edge segments for the invariant features. As we describe in Chapter 1, our final goal is generating robust background model and segmenting moving objects when camera has 3-Dimensional transformation in real time. Hence, we need to know the direction of feature and arranging order by importance. The SIFT algorithm [7] shows how we can obtain the direction of feature from

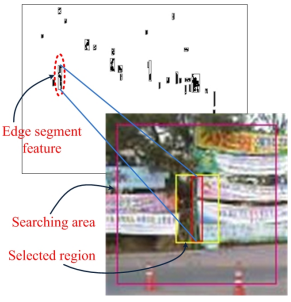


Fig. 2. Description of selected region with edge segment feature and searching area

corner point during descriptor step. However it is impossible to arrange corner features by importance while edge segments are possible. The edge segment features are selected from gray scale image. The Sobel operator is used to calculate each components of vertical and horizontal edges from gray scale image. However horizontal or vertical edge components are still remained in their opposite components when an edge is slanted. Therefore vertical and horizontal edges are acquired by Eq.(1).

$$\begin{aligned} I_V^{'} &= S_V * I - S_H * I \\ I_H^{'} &= S_H * I - S_V * I \end{aligned} \tag{1}$$

In the Eq.(1), S_V, S_H are vertical and horizontal Sobel kernels respectively, I is original image. Moreover $I_V^{'}$ and $I_H^{'}$ are vertical and horizontal edge images respectively. However, if variation of a scene is exactly same as camera motion, it is impossible to calculate camera motion nevertheless we have edge components. Therefore remained edge segments are derived by Eq.(2)

$$I_F = \left(I_V^{'} \bigcap I_{diff}\right) \bigcup \left(I_H^{'} \bigcap I_{diff}\right) \tag{2}$$

I_{diff} means temporal difference image and I_F is image of edge segment features in Eq.(2). The Fig.1 shows the selected edge segment features. The Fig.1(a) is original image and Fig.1(e) is temporal difference. Fig.1(b),(f) are vertical and horizontal edges and Fig.1(c),(g) are edge segment features. Using edge segment features, we choose 'selected region' that includes each edge segment features in current image. Then we choose 'searching area' shown in Finally, Fig.1(d),(h) and it is clearly described in Fig.2. The 'searching area' includes each edge segment features and belongs in previous image to calculate the similarity between two consecutive images.

2.2 Camera Motion Vector

In Eq.(3), k is the number of searching windows in the searching area. It is dependent on size of searching area and selected region. Between searching window

k and selected region, we calculate similarity ρ_t using correlation described in Eq.(3). (μ_t^s, σ_t^s) and (μ_{t-1}^k,σ_{t-1}^k) are mean and standard deviation for selected region in current image and searching window in prior image. After we calculate similarity for whole searching area, we select best matched region by Eq.(4).

$$\begin{aligned}\rho_t(k) &= \frac{cov(I_t(s), I_{t-1}(k))}{\sigma_t^s \cdot \sigma_{t-1}^k} \\ &= \frac{E\left((I_t(s) - \mu_t^s)(I_{t-1}(k) - \mu_{t-1}^k)\right)}{\sigma_t^s \cdot \sigma_{t-1}^k}\end{aligned} \tag{3}$$

$k = 1, \cdots, m$
$m = (w_p - w_c + 1) \times (h_p - h_c + 1)$
$I_t(s)$: pixels in the selected region
$I_{t-1}(k)$: pixels in the searching window
w_p, h_p : width and height of searching area
w_c, h_c : width and height of selected region

$$k_t = \arg\max_k \left(\rho_t\left(k\right) > 0.9\right) \tag{4}$$

When we find the best matched region k_t, we easily compose displacement vector based on Eq.(5).

$$\begin{aligned}|\upsilon_k| &= \sqrt{(i_p - i_c)^2 + (j_p - j_c)^2} \\ \theta_k &= \tan^{-1}\frac{j_p - j_c}{i_p - i_c}\end{aligned} \tag{5}$$

υ_k and θ_k in Eq.(5) are magnitude and its angle for displacement vector described the movement of each best matched region. (i_c, j_c) and (i_p, j_p) are center pixels in current and prior image respectively. Among displacement vectors, we choose a camera motion vector which has high density compare with others.

3 Multiple Background Model

3.1 Classification of Clusters for Multiple Background Model

In this paper, we protect to be blended background model by multiple clusters for each pixel. When we have best matched region by correlation coefficient, we classify clusters for each pixel based on camera flow. It is described below.

Initial value of the δ is 10

i. Initialization: $C^i(p) = I_t(p), \quad M^i(p) = 1 \quad i = 1$
ii. Finding nearest cluster with current pixel $I_t(p)$ by their absolute difference.

iii. If the absolute difference between $I_t(p)$ with nearest cluster $C^i(p)$ is smaller than threshold δ, updating number and center of cluster.

$$M^i(p) = M^i(p) + 1$$
$$C^i(p) = \left(I_i(p) + \left(M^i(p) - 1\right) * C^i(p)\right) / M^i(p)$$

iv. Otherwise, new cluster will be generated.

$$C^{i+1}(p) = I_t(p)$$
$$M^{i+1}(p) = 1$$

v. Iterating steps from (ii) to (iv) until no more pattern remained.

Notice. $C^i(p)$ and $M^i(p)$ are mean and member of i^{th} cluster in a pixel p respectively.

The next step is classification of clusters and eliminating small clusters which describe unexpectedly included clusters by moving objects. It is also described below.

---

i. Calculating weight for each cluster

$$W^i(p) = \frac{M^i(p)}{\sum^i M^i(p)} \quad i = 1, 2, \cdots, G(p) \tag{6}$$

ii. Eliminating small clusters among $G(p)$ depend on $W^i(p)$. Probabilistic threshold ζ (it is 0.2 in this paper chosen by experiment) is used in this process.

iii. Generating MBM with their weight

$$B^j(p) = C^j(p), \quad j = 1, 2, \cdots, S(p) \tag{7}$$

$$W^j(p) = \frac{M^j(p)}{\sum^j M^j(p)} \tag{8}$$

Notice. $W^i(p)$ and $B^i(p)$ are weight and MBM of i^{th} cluster in a pixel p respectively.

3.2 Updating Background

A background is normally and temporally changed. Therefore background update is important problem not only detection of moving object but also understanding environment in time sequence. One of simple classification of a background is

long-term and short-term background discussed by Elías et al. [5]. In their updating strategy, they use temporal median filter to generate initial background and update background using stable information in long time (long-term background) and temporal changes (short-term background). For our case, the progress of generating multiple background concerns classifying the clusters of stable information. However, when a car will park in the scene or leave, it is difficult to arrange the cluster that has previous information. For this reason, we use temporal median filter to eliminate the effect of non-exist cluster.

4 Experiments

In experiment, Firstly we experiment the proposed algorithm to the case of static camera. It is described in [12]. However our purpose is estimation of MBM and detection of moving objects by moving or shaking camera. In this case, MBM easily have the effect of temporal changes. Therefore we use temporal median filter to overcome it. Examples of edge segment features from moving camera are

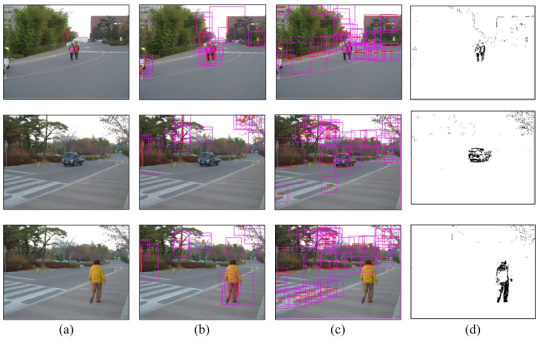


Fig. 3. Edge segment features from the sequence of moving camera

shown in Fig.3. Fig.3 shows two image sequences. The first row is one sequence and second and third rows are the other sequence but different time periods. Furthermore Fig.3(a) is original images and Fig.3(b) and Fig.3(c) are vertical and horizontal edge segment respectively with searching areas. From the image sequences by moving camera, we estimate MBM with temporal median background model shown in Fig.4. From Fig.4(a) to Fig.4(d) are 1^{st} to 4^{th} MBM and Fig.4(e) is temporal-median background. Temporal-median background is generated using median value for recent 7 frames. Using the MBM, we extract the moving objects shown in Fig.3(d). Even though it still have small noises in the Fig.3(d), most parts of moving objects are successfully detected. However some parts of the moving object are rejected. For example, the couple in first row and the woman in the last row in Fig.3 wear pink, red and yellow jumpers respectively. Those are distinctive colors from background, however those are

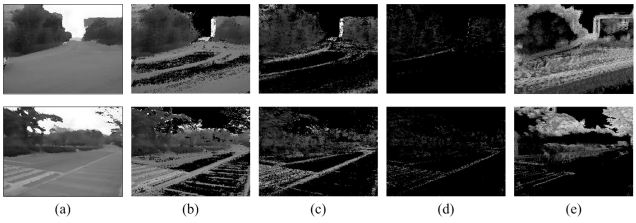


Fig. 4. MBM and temporal-median background for moving camera

Table 1. Processing time for each step

Process	Edge seg.	Correlation	Background model	Object detection	Tot
Time(ms)	3	11	7	1	22

rejected in the result. These errors are occurred by converting RGB color model to gray scale. Nevertheless those people wear significant colors, their converted value in gray scale is almost similar to their background model. With all its minor defects, we still have moving object shown in Fig.3(d). Therefore it is no exaggeration to say that our proposed algorithm is strongly detect moving objects while the background model is generated simultaneously.

5 Conclusions

We propose the algorithm of generating MBM(Multiple Background Model) and detection of moving objects under camera movement in this paper. In the section 2.1, we use feature based similarity calculation. Because we use strong edge segments to the features, calculated result shows low computational cost and similar performance of calculating camera motion by searching whole image. Furthermore, as we show in Chapter 4, we effectively estimate MBM by estimated camera motion using edge segment features. However MBM has weakness when a background is temporally changed. To overcome this problem, we use temporal-median filter. The result of detecting moving objects in the Fig.3(d) shows our method is good enough to detect moving object even though camera moves. However, our approach has limitation of camera motion described in Chapter 1. In this paper, we assume that camera has only 2-Dimensional translation. Other two categories in Chapter 1 and the case of fusing three categories are still researched. In spite of those limitation, proposed algorithm successfully generate multiple background model and detect moving object with low computational cost shown in Table 1. As the experimental results, we show that our proposed algorithm handles to generate stable background using not only stationary but also moving or shaking camera.

Acknowledgment

The authors would like to thank to Ulsan Metropolitan City and MOCIE and MOE of Korean Government which partly supported this research through the NARC and post BK21 project at University of Ulsan.

References

1. Wren, C.R., Azarbayejani, A., Darrell, T., Pentland, A.P.: Pfinder:Real-time tracking of the human body. IEEE Tran. on Pattern Analysis and Machine Intelligence 19(7), 780–785 (1997)
2. Kornprobst, P., Deriche, R., Aubert, G.: A real time system for detection and tracking people. Int. J. of Mathematical Imaging and Vision 11(1), 5–26 (1999)
3. Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Proc. of Int. Conf. on Computer Vision and Pattern Recognition, pp. 246–252 (1999)
4. Elgammal, A., Duraiswami, R., Harwood, D., Davis, L.: Background and foreground modeling using nonparametric kernel density estimation for visual Surveillance. Proc. of the IEEE 90(7), 1151–1163 (2002)
5. Elías, H.J., Carlos, O.U., Jesús, S.: Detected motion classification with a double-background and a Neighborhood-based difference. Pattern Recognition Letters 24(12), 2079–2092 (2003)
6. Mei, H., Kanade, T.: Multiple motion scene reconstruction with uncalibrated cameras. IEEE Tran. on Pattern Analysis and Machine Intelligence 25(7), 884–894 (2003)
7. David, G.L.: Distinctive Image Features from Scale-Invariant Keypoints. Int. J. of Computer Vision 2(60), 91–110 (2004)
8. Kang, J., Cohen, I., Medioni, G.: Tracking objects from multiple and moving cameras. IEE Intelligent Distributed Surveillance Systems, 31–35 (2004)
9. Li, L., Huang, W., Gu, I.Y.H., Tian, Q.: Statistical modeling of complex backgrounds for foreground object detection. IEEE Tran. on Image Processing 13(11), 1459–1472 (2004)
10. Xiao, M., Han, C., Kang, X.: A background reconstruction for dynamic scenes. In: Proc. of Int. Conf. on Information Fusion, pp. 1–7 (2006)
11. Chang, Y., Medioni, G., Jinman, K., Cohen, I.: Detecting Motion Regions in the Presence of a Strong Parallax from a Moving Camera by Multiview Geometric Constraints. IEEE Tran. on Pattern Analysis and Machine Intelligence 29(9), 1627–1641 (2007)
12. Taeho, K., Kang-Hyun, J.: Detection of Moving Object Using Remained Background under Moving Camera. Int. J. of Information Acquisition 4(3), 227–236 (2007)

Cross Ratio-Based Refinement of Local Features for Building Recognition

Hoang-Hon Trinh, Dae-Nyeon Kim, and Kang-Hyun Jo

Graduate School of Electrical Engineering, University of Ulsan, Korea
San 29, Mugeo-Dong, Nam-Ku, Ulsan 680 - 749, Korea
{hhtrinh,dnkim2005,jkh2008}@islab.ulsan.ac.kr

Abstract. This paper describes an approach to recognize buildings. The characters of building such as facets, their area, vanishing points, wall histogram and a list of local features are extracted and then stored in a database. Given a new image, the facet with biggest area is compared against the database to choose the closest pose. Novel methods of cross ratio-based refinement and SVD (singular value decomposition) based method are used to increase the recognition rate, increase the number of correspondences between image pairs and decrease the size of database. The proposed approach has been performed with 50 interest buildings containing 1050 images and a set of 50 test images. All images are taken under general conditions like different weather, seasons, scale, viewpoints and multiple buildings. We obtained 100(%) recognition rate.

Keywords: Cross ratio, wall histogram, SVD, multiple buildings.

1 Introduction

Many methods were proposed to solve problems of object recognition. Among them, three methods are widely known [6] such as appearance-based method [2], geometry-based method [1] and local feature-based method [4,7]. The local feature-based method is widely used because it is generality, robustness and easy learning. A local feature is represented by a detector (keypoint) and a descriptor. The detector should be frequently appeared and localized in different poses. The color information of the surrounding region of detector is used to encode in a feature vector. A major limitation is that many descriptors are stored. One object is appeared several poses (e.g., 5 poses in [7]). For each pose, hundreds of descriptors stored. So the number of mismatches increases and affects the recognition rate. The closest pose is chosen by the maximum number of total matches including mismatches and correct matches [7] or just only the correct matches [4]. The correct matches are verified from total matches by geometrical relation between the image pairs. The verification is usually performed by RANSAC (random sample Consensus) [5] or Hough transform [4] method. The geometrical transformation-based RANSAC is usually used when the inliers are larger than 50 percent. Whereas, when the percentage of inliers is small, the affine transformation-based Hough transform is used [4]. In practice, building

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 544–551, 2008.
© Springer-Verlag Berlin Heidelberg 2008

is a large planar objects so the mentioned methods are not preeminent. In the general case, the mismatches can not be used for selecting the closest pose. But the building is a repeated structure object [7]. So a part of mismatches can be useful for decision of the closest pose.

In this paper, the local feature-based recognition is used within two steps. The first step is for selecting a sub-candidate. Given a new image, we first detected the facets and their characters of building. Then the wall histogram and area are used to match the database for selecting a small number of candidates. The second step is for refining the closest pose by using the local features. We concern about several problems which directly effect to the recognition rate such as reducing the size of database, increasing the number of correct matches between the image pairs. To do so, the database is constructed by only one model for each building. The model then is updated by a training set. For updating database, the correspondences of image pairs must be exactly verified. Two novel methods, cross ratio-based refinement and SVD-based update, are proposed for verifying the correct matches and updating the model, respectively.

2 Characters of Building Image

The building is represented by the number of facets. Each facet is characterized by an area, wall histogram, vertical and horizontal dominant vanishing points (DVPs), and a list of local features. The process of facet detection was explained in detail in our previous works [8,9]. We first detected line segments and then roughly rejected the segments which come from the scene as tree, bush, sky and so on. MSAC (m-estimator sample consensus) algorithm is used for clustering segments into the common DVPs comprising one vertical and several horizontal vanishing points. The number of intersections between the vertical line and horizontal segments is counted to separate the building pattern into the independent facet. Finally, the boundaries of facet were found. Fig. 1 illustrates some examples of facet detection. The first row are the results of detected multiple buildings. The different colors represent the faces with different horizontal vanishing points. The second row shows that the images are taken under the general conditions.

Fig. 1. The results of facet detection

2.1 Wall Histogram

For the first step of match, several previous works used hue color information [7,8,9]. A major drawback of hue color histogram is not highly identifiable for gray level images. We overcome this problem by a new histogram. Firstly, a hue histogram of facet's pixels is calculated as the dark graph of Fig. 2(b). Then it is smoothed several times by 1D Gaussian filter. The peaks in the smoothed histogram are detected, and then the continuous bins that are larger than 40(%) of the highest peak are clustered into separate groups. In Fig. 2(b), the light graph is smoothed histogram, we obtain two separate groups. The pixels indexed by each continuous bin group are clustered together. Figs. 2(c,d) show two pixel groups. Each pixel group is segemnted again where the hue values are is replaced by gray intensity information. Finally, the biggest group of pixels is chosen as wall region as in Fig. 2(e). The information of wall pixels only is used to encode a 36-bin histogram with three components h_1, h_2 and h_3. h_1 (h_2) is a 10-bin histogram of H_1 (H_2) which is calculated by

$$H_1 = \arctan(\frac{R}{\alpha G}); H_2 = \arctan(\frac{G}{\beta B}) \Rightarrow 0 \leq H_1, H_2 \leq \frac{\pi}{2} \quad (1)$$

where, $\alpha(= 1.1)$ and $\beta(= 1.05)$ are compensation factors because the change of RGB components is different under the sunlight. $h_3 = \bar{s}h_{hue}$, where h_{hue} is a 16-bin histogram of hue value; and $\bar{s}$ is the average of the saturation in the HSV (hue, saturation, value) color space. h_1, h_2 and h_{hue} are normalized to unit length. Finally, the wall histogram is created by concatenating h_1, h_2 and h_3 and then again normalized. With gray color like Fig. 2(a), the value $\bar{s}$ is small, so it reduces the affection of hue component (Fig. 2(f)). Whereas, for the more specific color, the value $\bar{s}$ is larger so it increases the effect of hue component like Figs. 2(i,l). Figs. 2(g-l) illustrate the robustness of wall histogram. Figs. 2(g,j) are two poses of a building under different illumination of sunlight. Figs. 2(h,k) are wall regions and Figs. 2(i,l) are correspondent wall histograms. The wall histograms are approximate together that their Chi-squared distance (χ^2) is 0.06.

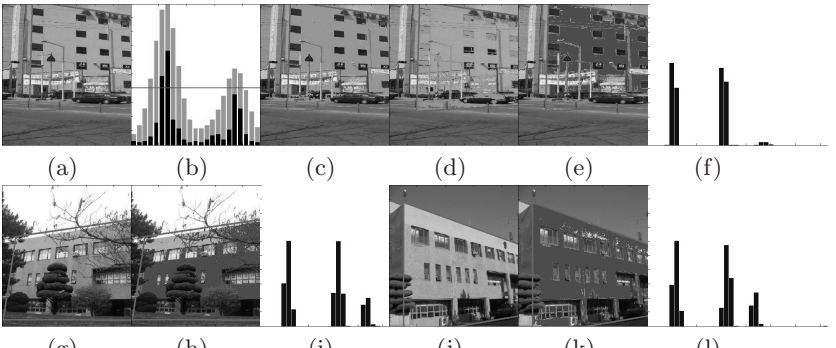


Fig. 2. The illustration of detection and robustness of wall histogram

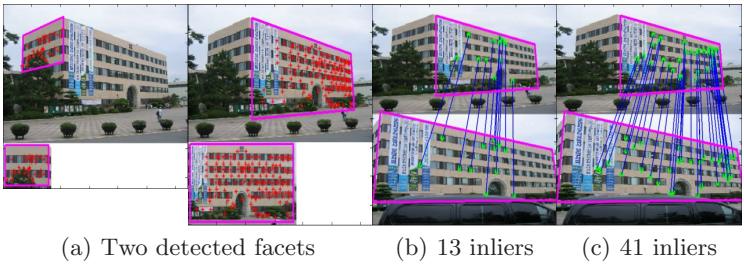
2.2 Local Features and Facet Transformation

To decrease the error of distortion, the facet is transformed into a rectangular shape which is similar to a frontal orthogonal view as on Fig. 3(a). The coordinates of 4 corners of facet's boundary are used to transform the skewed parallelogram to the rectangle with the transformation matrix H_r. The width (height) of rectangle equals the average lengthes of horizontal (vertical) boundaries of the skewed parallelogram. Now, the distortion of facets is just affected by the scale and stretched transformation, while the rotated affection is remarkably decreased. We used SIFT (scale-invariant feature transform) descriptors [4] to represent the local features. We select that the keypoint's number is proportional to facet area. With 640x480 image size and 700 keypoints for the maximum size of facet, the number of keypoints in each facet is calculated by $N = 700 \times \frac{S}{640 \times 480}$, where S is area of facet. If there are more than N keypoints then just N largest scale keypoints are selected as Fig. 3(a).

The effect of using transformed facets is demonstrated in Figs. 3(b,c). There are two images with large range of scale and rotation. They are directly matched together with threshold-based constraint. Two descriptors are matched if their χ^2 is less than 2. Fig. 3(b) is 13 obtained correspondences when the local features are calculated by the original images. Fig. 3(c) shows 41 correct matches when the descriptors are computed on the transformed facets. Then the green marks on the Fig. 3(c) are estimated by the corresponding matrix H_r. Here, the correspondences increase about three times when we use the transformed facet.

3 A New Method for Verification of Correct Matches

To more exactly verify the correspondences of the image pairs of large planar object, A new method is proposed by cross ratio invariance. The cross-ratio $\rho(ABCD)$ of four collinear points A, B, C, D is defined as the "double ratio", $\rho(ABCD) = \frac{CA}{CB} : \frac{DA}{DB}$. If two segments intersect four concurrent lines at A, B, C, D and A', B', C', D' like Fig. 4(a), their cross ratios is satisfied $\rho(ABCD) = \rho(A'B'C'D')$ [5]. Now, we consider a planar object as Fig. 4(b)



(a) Two detected facets (b) 13 inliers (c) 41 inliers

Fig. 3. (a) Two detected and corresponding transformed facets; detected keypoints (red marks); (b,c) correct matches with the original and transformed images, respectively

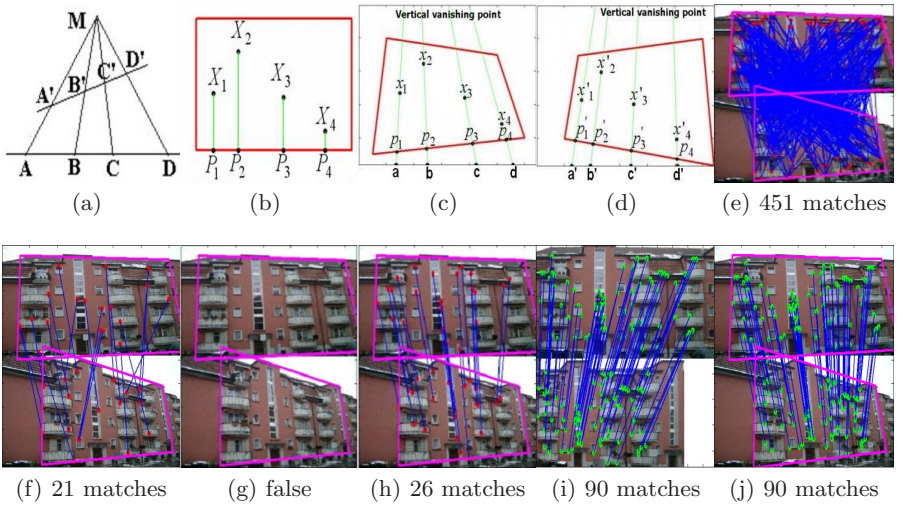
with four interest points, $\{X_1, X_2, X_3, X_4\}$, and rectangular boundary. Let $\{P_i\}$ points be the projections of $\{X_i\}, (i = 1, ..., 4)$ on the bottom boundary. Therefore, four lines P_iX_i parallel together. Assume that Figs. 4(c, d) are two different poses of this object where $\{x_i\}$ and $\{x'_i\}$ are the images of $\{X_i\}$. Similarly, $\{p_i\}$ and $\{p'_i\}$ are the images of $\{P_i\}$. In Fig. 4(c), Four lines $\{p_ix_i\}$ are concurrent at a vertical vanishing point. Let a, b, c, d be the intersections of $\{p_ix_i\}$ and the x-axis. Two set of points $\{a, b, c, d\}$ and $\{p_i\}$ are satisfied above property. Therefore, $\rho(abcd) = \rho(p_1p_2p_3p_4)$. Similarly, for Fig. 4(d), $\rho(a'b'c'd') = \rho(p'_1p'_2p'_3p'_4)$.

On the other hand, the sets of $\{p_i\}$ and $\{p'_i\}$ are projections of four collinear points $\{P_i\}$. So their cross ratios are invariant [5]. Finally, we have

$$\frac{(x_c - x_a)(x_d - x_b)}{(x_d - x_a)(x_c - x_b)} = \frac{(x_{c'} - x_{a'})(x_{d'} - x_{b'})}{(x_{d'} - x_{a'})(x_{c'} - x_{b'})} \tag{2}$$

Note that, if $x_a > x_b > x_c > x_d$ then $x_{a'} > x_{b'} > x_{c'} > x_{d'}$. This order is considered as a constraint in our method. Given two planar images with available vanishing points and N correspondences $\{X_i \longleftrightarrow X'_i\}$, $i = 1, 2, ..., N$; Let $\{x_i\}$ and $\{x_{i'}\}$ are projections of the correspondences on the x-axis of each image through the corresponding vanishing points, respectively. Randomly choosing a subset of three correspondences $\{a, b, c\}$ and $\{a', b', c'\}$, the error of cross ratio of the i^{th} correspondence following x-axis is defined

$$e_i^x = \frac{(x_i - x_a)(x_c - x_b)}{(x_c - x_a)(x_i - x_b)} - \frac{(x_{i'} - x_{a'})(x_{c'} - x_{b'})}{(x_{c'} - x_{a'})(x_{i'} - x_{b'})} \tag{3}$$



(a) (b) (c) (d) (e) 451 matches

(f) 21 matches (g) false (h) 26 matches (i) 90 matches (j) 90 matches

Fig. 4. (a) Cross ratio of four concurrent lines; (b) planar object; (c, d) two different poses; (e-j) the results of an example

Similar x-axis, on the y-axis, we get the error e_i^y.

Finally, the verification of correct matches is solved by Eq. 4 with RANSAC method.

$$minimize \sum((e_i^x)^2 + (e_i^y)^2); \quad i = 1, 2, ..., N \tag{4}$$

Here, the 2D problem is transformed into 1D problem. Fig. 4(e) is two far different point of views from ZuBud data [3] with 451 matches when the original images are used to calculate the local features. Fig. 4(f) containing 21 inliers is given from affine transformation and Hough transform [4]; there are several mismatches. Fig. 4(g) is false result when we used the general 2D projection based RANSAC method. Fig. 4(h) is the result of our proposed method. We obtaned 26 inliers. Here, the false correspondences are strongly rejected so we got the good result even in the case of less than 6% inliers. Figs. 4(ij) are the results of refinement by using the transformed facets. There are 90 correct matches with three times larger than when the original facets were used.

4 Model Update by Using SVD-Based Method

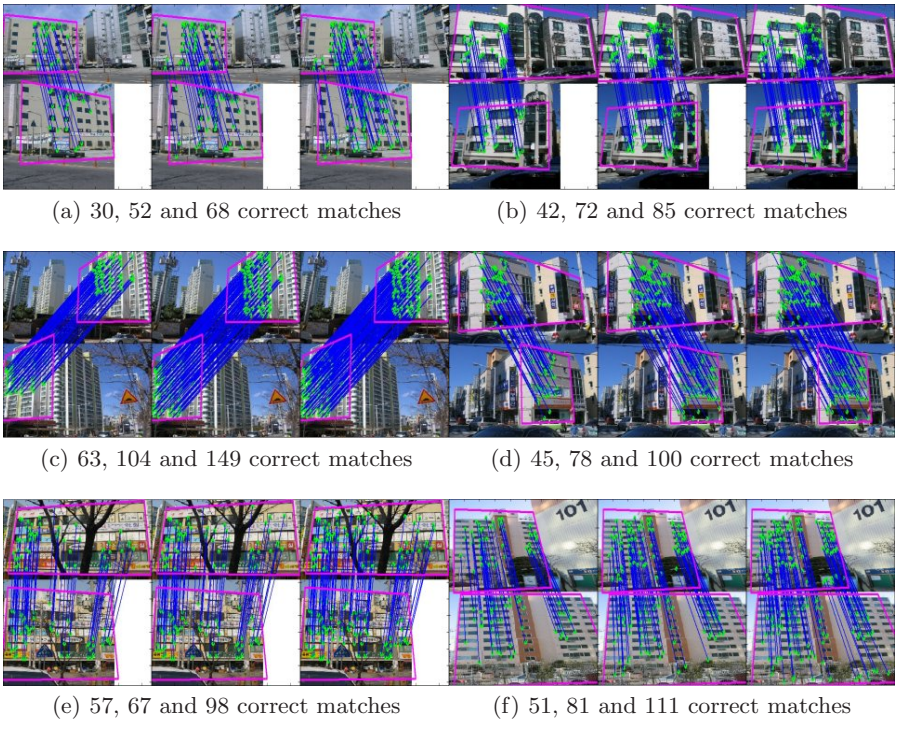
Given $n \times 2$ matrix A, we use the singular value decomposition algorithm to decompose the matrix A. $A = U \sum V^T$, where $\sum = diag(\lambda_1,\ \lambda_2)$. Let $a_1,\ a_2$ be the columns of A, if $\chi^2_{(a_1,a_2)} \simeq 0$ then $\lambda_1 \gg \lambda_2$ and $\lambda_2 \simeq 0$. An approximate matrix A' is calculated $U \sum' V^T$ where $\sum' = diag(\lambda_1, 0)$. Two columns $a'_1,\ a'_2$ of A' equal together after normalizing to unit length. Let $a = a'_1$, a is called approximate vector of column a_1 and a_2. For training stage, the database is updated step-by-step by supervised training strategy as follows,

1. *Initial model*: one image is taken for each building and then the characters of transformed facets are calculated and stored in a database.
2. *The first update*: Given a training image (T), just only the biggest area facet is chosen for matching to the database.
   Assume that a model M^i of database is matched to T, the correspondences of matched facet of M^i are replaced by the approximate vectors. Then the remained facets of T are only compared against to the remained ones of M^i. If the matching is successful then the matched facet is updated, otherwise that facet is added into M^i as a new facet. Finally, The T is stored at another place as an auxiliary model M^a with indexing to the M^i.
3. *The next update*:
   Firstly, the correct matches are updated.
   Secondly, the correspondences of T and M^a are calculated then updated to the M^i by corresponding transformative matrices. The M^a is replaced by T. If the number of keypoints in the facet is larger N then some keypoints whose updated times smallest are ruled out.

The updated model is affected by the noise from the training one. To overcome this problem, we decompose matrix A with a control factor. $A = [\gamma a_1, a_2]$, where $\gamma(= 2)$ is a control factor. Similarly, the color histogram is updated by the approximate vector.

5 Experiments

The proposed method has been performed with the real urban images contained 50 interest buildings and their neighborhoods. 22 poses are taken under general conditions for each building; The distribution is 1, 20 and 1 image for database, training set and test set, respectively. Given a test facet, the recognition progress comprises two stages. Firstly, by using a ratio area and χ^2 distance of wall histogram, a small set of candidates is chosen. The thresholds of the ratio area and histogram distances are fixed by [1/2, 2] and 0.1, respectively. Secondly, the recognition is refined by matching the SIFT features. To preserve the repeated structure of building, the threshold-base constraint is used; a test feature can have several matches whose distances are satisfied $d \leq 1.25 d_{smallest}$. Just only the biggest area facet of test image is considered for matching. Here, 78 facets are detected from database. The average area of detected facet is 52.73(%) the size of image. So each image contains about 369 keypoints. That size is 6.775 times smaller than the database's size of the approach of W. Zhang [7]. That approach stores 5 poses per building and around 500 keypoints for each pose in database. The results show that the number of correspondences increases about



(a) 30, 52 and 68 correct matches (b) 42, 72 and 85 correct matches

(c) 63, 104 and 149 correct matches (d) 45, 78 and 100 correct matches

(e) 57, 67 and 98 correct matches (f) 51, 81 and 111 correct matches

Fig. 5. From left to right, The results without, 10 and 20 time updating, respectively

50(%) after ten times of update. Fig. 5 shows several examples. For each sub-image, the above (bottom) image is the test (stored) building. In Figs. 5(a-f), the first, middle and the last images are the obtained correspondences of without, ten and twenty time updating, respectively. We obtain 100(%) recognition rate for 50 test images and the average size of sub-candidate is 9.5 facets.

6 Conclusions

A novel method is proposed for increasing the correspondences of local features; decreasing the size of database. Firstly, the building facets are detected and transformed into the rectangular shape for reducing the distortion error. The characters and SIFT descriptors are calculated by the transformed facets. Secondly, the cross ratio is used for correspondence verification that is high accuracy with large planar objects. Here, 2D problem is replaced by 1D problem that reduces the number of RANSAC loop. Finally, the SVD-based method is used for updating the database. So the size of database is decrease. The proposed method has performed to recognize 50 buildings with 100% recognition rate after 20 time updating. We are currently investing to apply the method for unsupervised training and larger database.

Acknowledgements

The authors would like to thank to Ulsan Metropolitan City and MOCIE and MOE of Korean Government which partly supported this research through the NARC and post BK21 project at University of Ulsan.

References

1. Pope, A.R.: Model-based object recognition: A Survey of Recent Research, Univ. of British Columbia (1994)
2. Swets, Weng, J.: Using discriminant eigenfeatures for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence 18(8), 831–836 (1996)
3. Shao, H., Gool, L.V.: Zubud-zurich Buildings Database for Image Based Recognition, Swiss FI of Tech., Tech. report no. 260 (2003)
4. Lowe, D.G.: Distinctive Image Features from Scale-invariant Keypoints. IJCV 60(2), 91–110 (2004)
5. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision, March, ch. 2, 4. Cambridge Uni. Press, Cambridge (2004)
6. Matas, J., Obdrzalek, S.: Object Recognition Methods based on Transformation Covariant Features. In: 12^{th} European Signal Processing Conf. (2004)
7. Zhang, W., Kosecka, J.: Hierarchical Building Recognition. IVC 25, 704–716 (2007)
8. Trinh, H.H., Kim, D.N., Jo, K.H.: Structure Analysis of Multiple Building for Mobile Robot Intelligence. In: SICE Proc., Japan (September 2007)
9. Trinh, H.H., Kim, D.N., Jo, K.H.: Urban Building Detection and Analysis by Visual and Geometrical Features. In: ICCAS 2007, Seoul, Korea (October 2007)

The Competitive EM Algorithm for Gaussian Mixtures with BYY Harmony Criterion

Hengyu Wang, Lei Li, and Jinwen Ma*

Department of Information Science, School of Mathematical Sciences and LAMA, Peking University, Beijing, 100871, China
jwma@math.pku.edu.cn

Abstract. Gaussian mixture has been widely used for data modeling and analysis and the EM algorithm is generally employed for its parameter learning. However, the EM algorithm may be trapped into a local maximum of the likelihood and even leads to a wrong result if the number of components is not appropriately set. Recently, the competitive EM (CEM) algorithm for Gaussian mixtures, a new kind of split-and-merge learning algorithm with certain competitive mechanism on estimated components of the EM algorithm, has been constructed to overcome these drawbacks. In this paper, we construct a new CEM algorithm through the Bayesian Ying-Yang (BYY) harmony stop criterion, instead of the previously used MML criterion. It is demonstrated by the simulation experiments that our proposed CEM algorithm outperforms the original one on both model selection and parameter estimation.

Keywords: Competitive EM (CEM) algorithm, Gaussian mixture, Bayesian Ying-Yang (BYY) harmony learning, Model selection

1 Introduction

As a powerful statistical tool, Gaussian mixture has been widely used in the fields of signal processing and pattern recognition. In fact, there already exist several statistical methods for the Gaussian mixture modeling, such as the k-means algorithm [2]) and the Expectation-Maximization (EM) algorithm [3]. However, the EM algorithm cannot determine the correct number of Gaussians in the mixture for a sample data set because the likelihood to be maximized is actually a increasing function of the number of Gaussians. Moreover, a "bad" initialization usually makes it trapped at a local maximum, and sometimes the EM algorithm converges to the boundary of the parameter space.

Conventionally, the methods of model selection for Gaussian mixture, i.e., determining a best number k^* of Gaussians for a sample data set, are based on certain selection criteria such as Akaike's Information Criterion (AIC) [4], Bayesian Information Criteria (BIC) [5], and Minimum Message Length (MML) criterion [6]. However, these criteria have certain limitations and often lead to a wrong result.

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 552–560, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Recently, with the development of the Bayesian Ying-Yang (BYY) harmony learning system and theory [7,8], there have emerged a new kind of learning algorithms [10,10,11,12] on the Gaussian mixture modeling based on the maximization of the harmony function, which is equivalent to the BYY harmony learning on certain architectures of the BYY learning system related to the Gaussian mixture model. These learning algorithms can automatically determine the number of Gaussians for the sample data set during parameter learning. Moreover, the successes of these algorithms also show that the harmony function can be served as an efficient criterion of model selection on Gaussian mixture.

Although these BYY harmony learning algorithms are quite efficient for the Gaussian mixture modeling, especially on automated model selection, they need an assumption that the number k of Gaussians in the mixture should be slightly larger than the true number k^* of Gaussians in the sample data. In fact, if k is smaller or too much larger than k^*, they may converge to a wrong result. On the other hand, it is still a difficult problem to estimate a reasonable upper bound of k^* with a sample data set. One possible way to overcome this difficulty is to introduce the split-and-merge operation or competitive mechanism into the EM algorithm to make the model selection dynamically [13]. However, the MML model selection criterion used in such a competitive EM (CEM) algorithm is not very efficient for the model selection on Gaussian mixture.

In this paper, we propose a new CEM algorithm which still works in a split-and-merge mode by using the BYY harmony criterion, i.e., the maximization of the harmony function, as a stop criterion. The simulation experiments demonstrate that our proposed CEM algorithm outperforms the original CEM algorithm on both model selection and parameter estimation.

2 The EM Algorithm for Gaussian Mixtures

We consider the following Gaussian mixture model:

$$p(x|\Theta_k) = \sum_{i=1}^{k} \alpha_i p(x|\theta_i) = \sum_{i=1}^{k} \alpha_i p(x|\mu_i, \Sigma_i), \tag{1}$$

where k is number of components in the mixture, $\alpha_i (\geq 0)$ are the mixing proportions of components satisfying $\sum_{i=1}^{k} \alpha_i = 1$ and each component density $p(x|\theta_i)$ is a Gaussian probability density function given by:

$$p(x|\theta_i) = p(x|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)}, \tag{2}$$

where μ_i is the mean vector and Σ_i is the covariance matrix which is assumed positive definite. For clarity, we let Θ_k be the collection of all the parameters in the mixture, i.e., $\Theta_k = (\theta_i, \cdots, \theta_k, \alpha_1, \cdots, \alpha_k)$.

Given a set of N i.i.d. samples, $\mathcal{X} = \{x_t\}_{t=1}^N$, the log-likelihood function for the Gaussian mixture model is expressed as follows:

$$\log p(\mathcal{X}|\Theta_k) = \log \prod_{t=1}^{N} p(x_t|\Theta_k) = \sum_{t=1}^{N} \log \sum_{i=1}^{k} \alpha_i p(x_t|\theta_i), \tag{3}$$

which can be maximized to get a Maximum Likelihood (ML) estimate of Θ_k via the following EM algorithm:

$$\alpha_i^+ = \frac{1}{n} \sum_{t=1}^{N} P(i|x_t); \mu_i^+ = \sum_{t=1}^{N} x_t P(i|x_t) / \sum_{t=1}^{N} P(i|x_t) \tag{4}$$

$$\Sigma_i^+ = [\sum_{t=1}^{N} P(i|x_t)(x_t - \mu_i^+)(x_t - \mu_i^+)^T] / \sum_{t=1}^{N} P(i|x_t), \tag{5}$$

where $P(i|x_t) = \alpha_i p(x_t|\theta_i) / \sum_{i=1}^{k} \alpha_i p(x_t|\theta_i)$ are the posterior probabilities.

Although the EM algorithm can have some good convergence properties, it clearly has no ability to determine the appropriate number of the components for a sample data set because it is based on the maximization of the likelihood function. In order to overcome this difficulty, we will use the BYY harmony function instead of the likelihood function for our Gaussian mixture learning.

3 BYY Learning System and Harmony Function

In a BYY learning system, each observation $x \in X \subset \mathcal{R}^d$ and its corresponding inner representation $y \in Y \subset \mathcal{R}^m$ are described with two types of Bayesian decomposition: $p(x, y) = p(x)p(y|x)$ and $q(x, y) = q(y)q(x|y)$, which are called them Yang and Ying machine, respectively. Given a data set $D_x = \{x_1, \cdots, x_n\}$ from the Yang or observation space, the task of learning on a BYY system consist of specifying all the aspect of $p(y|x)$,$p(x)$,$q(x|y)$,$q(y)$ with a harmony learning principle implemented by maximizing the function:

$$H(p \parallel q) = \int p(y|x)p(x) \ln[q(x|y)q(y)]dxdy. \tag{6}$$

For the Gaussian mixture model, we let y be limited to be an integer variable $y = \{1, \cdots, k\} \subset R$ and utilize the following specific BI-Architecture of the BYY system:

$$p(x) = p_0(x) = \frac{1}{N} \sum_{t=1}^{N} G(x - x_t); \qquad p(y = i|x) = \alpha_i q(x|\theta_i)/q(x|\Theta_k); \tag{7}$$

$$q(x|\Theta_k) = \sum_{i=1}^{k} \alpha_i q(x|\theta_i); \quad q(y) = q(y = i) = \alpha_i > 0; \quad \sum_{i=1}^{k} \alpha_i = 1, \tag{8}$$

where $G(\cdot)$ is a kernel function and $q(x|y=i) = q(x|\theta_i)$ is a Gaussian density function. In this architecture, $q(y)$ is a free probability function, $q(x|y)$ is a Gaussian density, and $p(y|x)$ is constructed from $q(y)$ and $q(x|y)$ under the Bayesian law. Θ_k=$\{\alpha_i, \theta_i\}_{i=1}^k$ is the collection of the parameters in this BYY learning system. Obviously $p(x|\Theta_k) = \sum_{i=1}^k \alpha_i q(x|\theta_i)$ is a Gaussian mixture model under the architecture of the BYY system.

Putting all these component densities into Eq.(6) and letting the kernel functions approach the delta functions, $H(p||q)$ reduces to the following harmony function:

$$H(p \parallel q) = J(\Theta_k) = \frac{1}{N} \sum_{t=1}^{N} \sum_{i=1}^{k} \frac{\alpha_i q(x_t|\theta_i)}{\sum_{j=1}^k \alpha_j q(x_t|\theta_j)} \ln[\alpha_i q(x_t|\theta_i)]. \tag{9}$$

Actually, it has been shown by the experiments and theoretical analysis that as this harmony function arrives at the global maximum, a number of Gaussians will match the actual Gaussians in the sample data, respectively, with the mixing proportions of the extra Gaussians attenuating to zero. Therefore, the maximization of the harmony function can be used as a reasonable criterion for model selection on Gaussian mixture.

4 The Competitive EM Algorithm with BYY Harmony Criterion

4.1 The Split and Merge Criteria

In order to overcome the weaknesses of the EM algorithm, we can introduce certain split-and-merge operation into the EM algorithm such that a competitive learning mechanism can be implemented on the estimated Gaussians obtained from the EM algorithm. We begin to introduce the split and merge criteria used in [13].

In order to do so, we define the local density $f_i(x;\Theta_k)$ as a modified empirical distribution given by:

$$f_i(x;\Theta_k) = \sum_{t=1}^{N} \delta(x - x_t) P(i|x_t;\Theta_k) / \sum_{t=1}^{N} P(i|x_t;\Theta_k), \tag{10}$$

where $\delta(\cdot)$ is the Kronecker function and $P(i|x_t;\Theta_k)$ is the posterior probability. Then, the local Kullback divergence can be used to measure the distance between the local data density $f_i(x;\Theta_k)$ and the estimated density $p(x|\theta_i)$ of the i-th component in the mixture:

$$D_i(\Theta_k) = \int f_i(x;\Theta_k) \log \frac{f_i(x;\Theta_k)}{p(x|\theta_i)} dx. \tag{11}$$

Thus, the split probability of the i-th component is assumed to be proportional to $D_i(\Theta_k)$:

$$P_{split}(i;\Theta_k) = (1/N_{\Theta_k}) \cdot D_i(\Theta_k), \tag{12}$$

where N_{Θ_k} is a regularization factor.

On the other hand, if the i-th and j-th components should be merged into a new component denoted as the i'-th component and the parameters of the new Gaussian mixture become Θ'_{k-1}, the merge probability of the i-th and j-th components is assumed to be inversely proportional to $D_{i'}(\Theta'_{k-1})$:

$$P_{merge}(i,k;\Theta_k) = (1/N_{\Theta_k}) \cdot (\beta / D_{i'}(\Theta'_{k-1})), \tag{13}$$

where β is also a regularization factor determined by experience. That is, as the split probability of the new component merged from the two old components becomes larger, the merge probability of these two components becomes smaller.

With the above split and merge probabilities, we can construct the following split and merge criteria.

Merge Criterion: If the i-th and j-th components gave the highest merge probability, we may merge them with the parameters θ_l of the new component or Gaussian as follows ([14]):

$$\alpha_l = \alpha_i + \alpha_j; \mu_l = (\alpha_i \mu_i + \alpha_j \mu_j)/\alpha_l; \tag{14}$$

$$\Sigma_l = \{\alpha_i[\Sigma_i + (\mu_i - \mu_l)(\mu_i - \mu_l)^T] \tag{15}$$

$$+\alpha_j[\Sigma_j + (\mu_j - \mu_l)(\mu_j - \mu_l)^T)]\}/\alpha_. \tag{16}$$

Split Criterion: If the r-th component has the highest split probability, we can split it into two components, called the i-th and j-th components. In order to do so, we get the singular value decomposition of the covariance matrix $\Sigma_r = USV^T$, where $S = diag[s_1, s_2, \cdots, s_d]$ is a diagonal matrix with nonnegative diagonal elements in a descent order, U and V are two (standard) orthogonal matrices. Then, we set $A = U\sqrt{S} = Udiag[\sqrt{s_1}, \sqrt{s_2}, \cdots, \sqrt{s_d}]$ and get the first column A_1 of A. Finally, we have the parameters for the two split components as follows ([14]):

$$\alpha_i = \gamma\alpha_r, \alpha_j = (1-\gamma)\alpha_r; \tag{17}$$

$$\mu_i = m_r - (\alpha_j/\alpha_i)^{1/2}\mu A_1, \mu_j = m_r + (\alpha_i/\alpha_j)^{1/2}\mu A_1; \tag{18}$$

$$\Sigma_i = (\alpha_j/\alpha_i)\Sigma_r + ((\eta - \eta\lambda^2 - 1)(\alpha_r/\alpha_i) + 1)A_1 A_1^T; \tag{19}$$

$$\Sigma_j = (\alpha_i/\alpha_j)\Sigma_r + ((\eta\lambda^2 - \eta - \lambda^2)(\alpha_r/\alpha_j) + 1)A_1 A_1^T, \tag{20}$$

where $\gamma, \mu, \lambda, \eta$ are all equal to 0.5. For clarity, we denote the parameters of the new mixture by Θ'_{k+1}.

4.2 The Proposed CEM Algorithm

The performance of the CEM algorithm strongly relies on the stop criterion. We now use the BYY harmony learning criterion as the stop criterion instead of the MML criterion in [6]. That is, our split-and merge EM learning process tries to maximize the harmony function $J(\Theta)$ given in Eq.(9). Specifically, in each iteration, if the harmony learning function $J(\Theta_k)$ increases, the split or merge operation will be accepted, otherwise it will be reject.

For quick convergence, during each stage of the algorithm, the component with a mixing proportion α_i being less than a threshold $\epsilon > 0$ will be discarded from the Gaussian mixture directly. With such a component annihilation mechanism, the algorithm can escape a lot of computation and converge to the solution more rapidly.

With all the preparations, we can summarize the CEM algorithm for Gaussian mixtures with the BYY harmony criterion as follows.

Step 1: Initialization. Select k and set the initial parameters Θ_k as randomly as possible. And set $l = 0$.

Step 2: Implement the (conventional) EM algorithm and get the new parameters as $\Theta_k(l)$ of the Gaussian mixture.

Step 3: Implement the Split and merge operations on the estimated components of the Gaussian mixture with $\Theta_k(l)$ and obtain the the new parameters $\Theta'_{k-1}(l)$ and $\Theta'_{k+1}(l)$ from the merge and split operations, respectively.

Step 4: Compute $Acc(M) = J(\Theta'_{k-1}(l)) - J(\Theta_k(l))$. If $Acc(M) > 0$, accept the merge operation, update the estimated mixture by $\Theta'_{k-1}(l)$, and go to Step 6. Otherwise, go to Step 5.

Step 5: Compute $Acc(S) = J(\Theta'_{k+1}(l)) - J(\Theta_k(l))$. If $Acc(S) > 0$, accept the split operation, update the estimated mixture by $\Theta'_{k+1}(l)$, and go to Step 6. Otherwise, stop the algorithm and get the parameters of the Gaussian mixture.

Step 6: Remove the components whose mixing proportions are lower than the pre-defined threshold $\epsilon > 0$. With the remaining parameters and k, let $l = l + 1$ and go to Step 2.

Clearly, this proposed CEM algorithm increases the harmony function on each stage and reach its maximum at end. The maximization of the harmony function as well as the EM algorithm guarantee the new CEM algorithm can lead to a good result on both model selection and parameter estimation for the Gaussian mixture modeling with a sample data set, which will be demonstrated by the simulation experiments in the next section.

5 Experimental Results

In this section, some simulation experiments were conducted to demonstrate the performance of the proposed CEM algorithm with BYY harmony criterion. Actually, the proposed CEM algorithm was implemented on two different synthetic data sets, each of which contained 3000 samples. Moreover, the proposed CEM algorithm was compared with the original CEM algorithm.

The first sample data set consisted of seven Guassians. As shown in the left subfigure of Fig.1, the proposed CEM algorithm was initialized with $k = 5$ from the parameters obtained by the k-means algorithm. After some iterations, as shown in the middle subfigure of Fig.1, one initial component was split into two Gaussians by the algorithm. Such a split process continued until all the

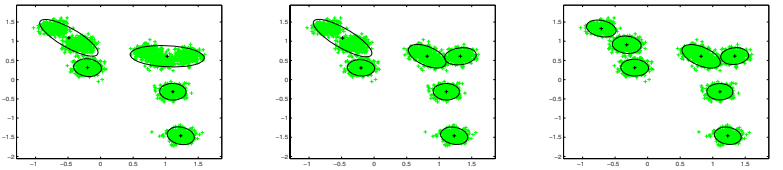


Fig. 1. The Experimental Results of the Proposed CEM Algorithm on the First Sample Data Set

components accurately matched the actual Gaussians in the sample data set, respectively, which is shown in the right subfigure of Fig.1. The algorithm stopped as $J(\Theta_k)$ reached its maximum and selected the correct number of Gaussians in the sample data set.

The second sample data set consisted of eight Gaussians. As shown in the left subfigure of Fig.2, the proposed CEM algorithm was implemented on the second sample data set initially with $k = 12$, which was much larger than the number of actual Gaussians. As shown in the middle subfigure of Fig.2, two pairs of the initially estimated Gaussians were selected to have been merged into two new Gaussians. The algorithm finally stopped with the eight actual Gaussians estimated accurately, being shown in the right subfigure of Fig.2.

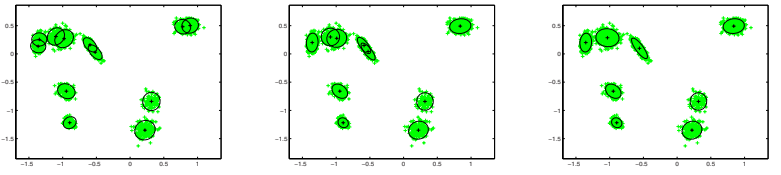


Fig. 2. The Experimental Results of the Proposed CEM Algorithm on the Second Sample Data Set

For comparison, we also implemented the original CEM algorithm with the MLL criterion [6] on the second sample data set. From Fig.3, it can be observed that the original CEM algorithm was initialized at the same situation, but led to

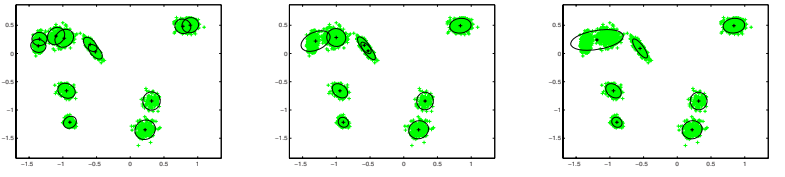


Fig. 3. The Experimental Results of the Original CEM Algorithm on the Second Sample Data Set

a wrong result on both model selection and parameter estimation at last. Therefore, our proposed CEM algorithm can outperform the original CEM algorithm on both model selection and parameter estimation in certain cases.

6 Conclusions

We have investigated the competitive EM algorithm from the view of the Bayesian Ying-Yang (BYY) harmony learning and proposed a new CEM algorithm with the BYY harmony criterion. The proposed competitive EM algorithm can automatically detect the correct number of Gaussians for a sample data set and obtain a good estimation of the parameters for the Gaussian mixture modeling through a series of the split and merge operations on the estimated Gaussians obtained from the EM algorithm. It is demonstrated well by the simulation experiments that the proposed CEM algorithm can achieve a better solution for the Gaussian mixture modeling on both model selection and parameter estimation on a sample data set.

Acknowledgements

This work was supported by the Natural Science Foundation of China for grant 60771061. The authors acknowledge Mr. Gang Chen for his helpful discussions.

References

1. Hartigan, J.A.: Distribution Problems in Clustering. Classification and Clustering. In: Van Ryzin, J. (ed.) Distribution Problems in Clustering, pp. 45–72. Academic press, New York (1977)
2. Jain, A.K., Dubes, R.C.: Algorithm for Clustering Data. Prentice Hall, Englewood Cliffs (1988)
3. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximun Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society B 39, 1–38 (1977)
4. Akaike, H.: A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control AC-19, 716–723 (1974)
5. Scharz, G.: Estimating the Dimension of a Model. The Annals of Statistics 6, 461–464 (1978)
6. Figueiredo, M.A.T., Jain, A.K.: Unsupervised Learning of Finite Mixture Models. IEEE Transations on Pattern Analysis and Machine Intelligence 24(3), 381–396 (2002)
7. Xu, L.: Best Harmony, Unified RPCL and Automated Model Selection for Unsupervised and Supervised Learning on Gaussian Mixtures, Three-Layer Nets and ME-RBF-SVM Models. International Journal of Neural Systems 11(1), 43–69 (2001)
8. Xu, L.: BYY Harmony Learning, Structural RPCL, and Topological Self-organizing on Mixture Modes. Neural Networks 15, 1231–1237 (2002)
9. Ma, J., Wang, T., Xu, L.: A Gradient BYY Harmony Learning Rule on Gaussian Mixture with Automated Model Selection. Neurocomputing 56, 481–487 (2004)

10. Ma, J., Wang, L.: BYY Harmony Learning on Finite Mixture: Adaptive Gradient Implementation and a Floating RPCL Mechanism. Neural Processing Letters 24(1), 19–40 (2006)
11. Ma, J., Liu, J.: The BYY Annealing Learning Algorithm for Gaussian Mixture with Automated Model Selection. Pattern Recognition 40, 2029–2037 (2007)
12. Ma, J., He, X.: A Fast Fixed-point BYY Harmony Learning Algorithm on Gaussian Mixture with Automated Model Selection. Pattern Recognition Letters 29(6), 701–711 (2008)
13. Zhang, B., Zhang, C., Yi, X.: Competitive EM Algorithm for Finite Mixture Models. Pattern Recognition 37, 131–144 (2004)
14. Zhang, Z., Chen, C., Sun, J., et al.: EM Algorithms for Gaussian Mixtures with Split-and-Merge Operation. Pattern Recogniton 36, 1973–1983 (2003)

Method of Face Recognition Based on Red-Black Wavelet Transform and PCA

Yuqing He, Huan He, and Hongying Yang

Department of Opto-Electronic Engineering,
Beijing Institute of Technology, Beijing, P.R. China, 100081
20701170@bit.edu.cn

Abstract. With the development of the man-machine interface and the recognition technology, face recognition has became one of the most important research aspects in the biological features recognition domain. Nowadays, PCA(Principal Components Analysis) has applied in recognition based on many face database and achieved good results. However, PCA has its limitations: the large volume of computing and the low distinction ability. In view of these limitations, this paper puts forward a face recognition method based on red-black wavelet transform and PCA. The improved histogram equalization is used to realize image preprocessing in order to compensate the illumination. Then, appling the red-black wavelet sub-band which contains the information of the original image to extract the feature and do matching. Comparing with the traditional methods, this one has better recognition rate and can reduce the computational complexity.

Keywords: Red-black wavelet transform, PCA, Face recognition, Improved histogram equalization.

1 Introduction

Because the traditional status recognition (ID card, password, etc) has some defects, the recognition technology based on biological features has become the focus of the research. Compared with the other biological features (such as fingerprints, DNA, palm prints, etc) recognition technology, people identify with the people around mostly using the biological characteristics of human face. Face is the most universal mode in human vision. The visual information reflected by human face in the exchange and contact of people has an important role and significance. Therefore, face recognition is the easiest way to be accepted in the identification field and becomes one of most potential identification authentication methods. Face recognition technology has the characteristics of convenient access, rich information. It has wide range of applications such as identification, driver's license and passport check, banking and customs control system, and other fields[1].

The main methods of face recognition technology can be summed up to three kinds: based on geometric features, template and model separately. The PCA face recognition method based on K-L transform has been concerned since the 1990s. It is simple, fast. and easy to use. It can reflect the person face's characteristic on the whole. Therefore, applying PCA method in the face recognition is unceasingly improving.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 561–568, 2008.
© Springer-Verlag Berlin Heidelberg 2008

This paper puts forward a method of face recognition based on Red-Black wavelet transform and PCA. Firstly, using the improved image histogram equalization[2] to do image preprocessing, eliminating the impact of the differences in light intensity. Secondly, using the Red-Black wavelet transform to withdraw the blue sub-band of the relative stable face image to obscure the impacts of expressions and postures. Then, using PCA to withdraw the feature component and do recognition. Comparing with the traditional PCA methods, this one can obviously reduce computational complexity and increase the recognition rate and anti-noise performance. The experimental results show that this method mentioned in this paper is more accurate and effective.

2 Red-Black Wavelet Transform

Lifting wavelet transform is an effective wavelet transform which developed rapidly these years. It discards the complex mathematical concepts and the telescopic and translation of the Fourier transform analysis in the classical wavelet transform. It develops from the thought of the classical wavelet transform multi-resolution analysis. Red-black wavelet transform[3-4] is a two-dimensional lifting wavelet transform[5-6], it contains horizontal/vertical lifing and diagonal lifting. The specific principles are as bellow.

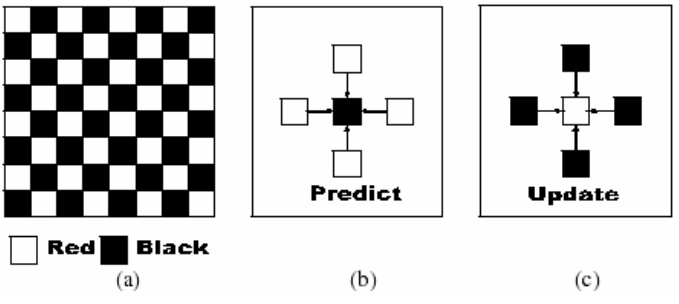


Fig. 1. Red-Black wavelet transform horizontal /vertical lifting

2.1 Horizontal /Vertical Lifting

As Fig.1 shows, horizontal /vertical lifting is divided into three steps:

1. Decomposition**:** The original image by horizontal and vertical direction is divided into red and black block in a cross-block way.

2. Prediction**:** Carry on the prediction using horizontal and the vertical direction four neighborhood's red blocks to obtain a black block predicted value. Then, using the difference of the black block actual value and the predicted value to substitute the black block actual value. Its result obtains the original image wavelet coefficient. As Fig.1(b) shows:

$$f(i,j) \leftarrow f(i,j) - \left[f(i-1,j) + f(i,j-1) + f(i,j+1) + f(i+1,j) \right] / 4$$
$$(i \bmod 2 \neq j \bmod 2) \quad (1)$$

3. Revision: Using the horizontal and vertical direction founeighborhood's black block's wavelet coefficient to revise the red block actual value to obtain the approximate signal. As Fig.1(c) shows:

$$f(i,j) \leftarrow f(i,j) + \left[f(i-1,j) + f(i,j-1) + f(i,j+1) + f(i+1,j) \right] / 8$$
$$(i \bmod 2 = j \bmod 2) \quad (2)$$

In this way, the red block corresponds to the approximating information of the image, and the black block corresponds to the details of the image.

2.2 Diagonal Lifting

On the basis of horizontal /vertical lifting, we do the diagonal lifting. As Fig.2 shows, it is also divided into three steps:

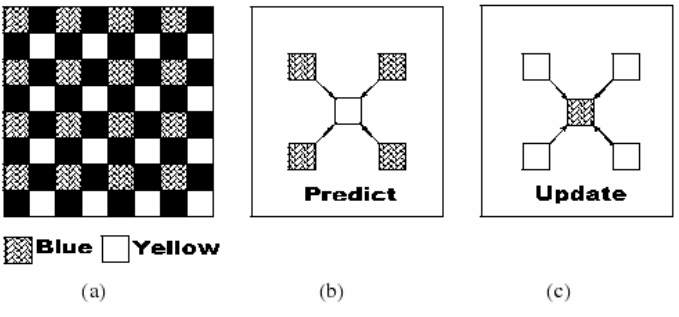


Fig. 2. Diagonal lifting

*1.*Decomposition: After horizontal /vertical lifting, dividing the obtained red block into the blue block and the yellow block in the diagonal cross way.

2. Prediction: Using four opposite angle neighborhood's blue block to predict a data in order to obtain the yellow block predicted value. Then the difference of the yellow block actual value and the predicted value substitutes the yellow block actual value. Its result obtains the original image wavelet coefficient of the diagonal direction. As Fig.2(b) shows:

$$f(i,j) \leftarrow f(i,j) - \left[f(i-1,j-1) + f(i-1,j+1) + f(i+1,j-1) + f(i+1,j+1) \right] / 4$$
$$(i \bmod 2 = 1, j \bmod 2 = 1) \quad (3)$$

3. Revision: Using four opposite angle neighborhood yellow block wavelet coefficient to revise the blue block actual value in order to obtain the approximate signal. As Fig.2(c) shows:

$$f(i,j) \leftarrow f(i,j) + \left[f(i-1,j-1) + f(i-1,j+1) + f(i+1,j-1) + f(i+1,j+1) \right] / 8 \quad (i \bmod 2 = 0, j \bmod 2 = 0) \quad (4)$$

After the second lifting, the red-black wavelet transform is realized.

According to the Equations, it can analyze some corresponding relations between the red-black wavelet transform and the classical wavelet transform: namely, the blue block is equal to the sub-band LL of the classical tensor product wavelets, the yellow block is equal to sub-band HH and the black block is equal to sub-band HL and LH. Experimental results show that it discards the complex mathematical concepts and equations. The relativity of image can mostly be eliminated and the sparser representation of image can be obtained by the Red-Black wavelet transform.

The image after Red-Black wavelet transform is showed in the Fig.3(b), on the left corner is the blue sub-band block image which is the approximate image of original image.

(a)Original image

(b)Image after Red-Black wavelet transform

Fig. 3. The result of red-black wavelet transform

3 Feature Extraction Based on PCA[7]

PCA is a method which analyses data in statistical way. This method discovers group of vectors in the data space. Using these vectors to express the data variance as far as possible. Putting the data from the P-dimensional space down to M-dimensional space (P>>M). PCA use K-L transform to obtain the minimum-dimensional image recognition space of the approximating image space. It views the face image as a high-dimensional vector. The high-dimensional vector is composed of each pixel. Then the high-dimensional information space maps the low-dimensional characteristic subspace by K-L transform. It obtains a group of orthogonal bases through high-dimensional face image space K-L transform. The partial retention of orthogonal bases creates the low-dimensional subspace. The orthogonal bases reserved is called

"Principle component". Since the image corresponding to the orthogonal bases just like face, so this is also called "Eigenfaces" method. The arithmetic of feature extraction are specified as follows:

For a face image of m × n, connecting its each row will constitute a row vector which has D= m × n dimensions. The D is the face image dimensions. Supposing M is the number of training samples, X_j is the face image vector which is derived from the jth picture, so the covariance matrix of the whole samples is:

$$S_T = \sum_{j=1}^{M} \left(X_j - \mu \right) \left(X_j - \mu \right)^T \tag{5}$$

And the μ is the average image vector of the training samples:

$$\mu = \frac{1}{M} \sum_{j=1}^{M} X_j \tag{6}$$

Ordering A=[X_1- μ , X_2− μ , ⋯, X_M− μ], so S_T=$AA^T$and its demision is D×D.

According to the principle of K-L transform, the coordinate we achieved is composed of eigenvector corresponding to nonzero eigenvalue of matrix AA^T. Computing out the eigenvalue and Orthogonal normalized vector of matrix D×D directly is difficult. So according to the SVD principle, it can figure out the eigenvalue and eigenvector of matrix AA^T through getting the eigenvalue and eigenvector of matrix A^TA.

λ_i （i=1,2,…,r） is r nonzero eigenvalue of matrix A^TA, v_i is the eigenvector corresponding to λ_i, so the orthogonal normalized eigenvector μ_i of matrix AA^T is as bellow:

$$\mu_i = \frac{1}{\sqrt{\lambda_i}} A v_i (i = 1, 2, ..., r) \tag{7}$$

This is the eigenvector of AA^T. Arranging its eigenvalues according to the size:$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_i$, its corresponding eigenvector is μ_i. In this way, each face image can project on the sub-space composed of μ_1, $\mu_2 \ldots \mu_r$. In order to reduce the dimension, it can select the former d eigenvectors as sub-space. It can select d biggest eigenvectors according to the energy proportion which the eigenvalue occupies:

$$\sum_{i=0}^{d} \lambda_i / \sum_{i=0}^{r} \lambda_i \geq \alpha \tag{8}$$

Usually ordering α =90%~99%.

As a result , the image corresponding to these eigenvectors are similar to the human face, it is also called "Eigenfaces".So the method which uses PCA transform is called"Eigenfaces"method. Owing to the Drop-dimensional space composed of "Eigenfaces", each image can project on it and get a group of coordinate coefficients which shows the location of the sub-space of this image, so it can be used as the bases for face recognition. Therefore, it can use the easiest Nearest Neighbor Classifier[8] to classify the faces.

4 Experiments and Results

This part mainly verifies the feasibility and superiority of the algorithm through the comparison of the experimental data.

4.1 Experimental Conditions and Parameters

Selecting Images After Delamination. After Red-Black wavelet transform, we only select the data of the blue block, because the part of the blue block represents the approximating face image and it is not sensitive to the expression and illumination and even filtrates the image noise.

The Influence of The Blue Block Energy Caused By Decomposition Layers. In the experiment data[9] of the Red-Black transform, we can find that it does not have the energy centralized characteristic under the situation of multi-layer decomposition. As the Table 1 shows that different layer decompositions obtain different energies. Test image is the international standard image Lena512, its energy is 4.63431e+009 and entropy is 7.4455.

Table 1. Red-Black wavelet energy decomposition test chart

Layers	1	2	3	4	5
Low-pass energy(%)	0. 9932	0. 9478	0. 7719	0. 4332	0. 1538
Total energy	1.17e+009	3.09e+008	9.58e+007	4.33e+007	3.06e+007

The original image energy wastage is due to the black and yellow block transform.

According to the former results and the size of the face image (112×92), one layer decomposition can be done to achieve satisfactory results. The blue block sub-band not only has no incentive to expression and gestures, but also retains the difference of different faces. At the same time it reduces the image vector dimensions and the complexity of the algorithm. If the size of the original image is bigger and the resolution is higher, the multi-layer decompositions can be considered.

Database Selection. We choose the public ORL database to do some related experiments. This database contains 40 different people's images which are captured in different periods and situation. It has 10 pictures per person and 400 pictures in all. The background is black. Each picture has 256 grayscales and the size is 112×92. The face images in the database have the different facial expressions and the different facial detail changes. The facial postures also have the changes. At present, this is the most extensive face database.

4.2 Experiment Processes and Results

The method of face recognition based on Red-Black wavelet transform and PCA shows as bellow: Firstly, using the improved image histogram equalization to do pretreatment,

Fig. 5. Original face images

Fig. 6. Images of illumination compensation results

Fig. 7. Images of one layer Red-Black wavelet transform results

eliminating the impact of the differences in light intensity. Secondly, using the Red-black wavelet transform to withdraw the blue block sub-band of the relative stable person face image achieved the effects of obscuring impacts of expressions and postures. Then, using PCA to withdraw the feature component and do recognition. We adopt 40 person and 5 pictures per person when training, so there are 200 pictures as the training samples all together. Then carrying on recognition to the other 200 pictures under the conditions of whether there are illumination and Red-Black wavelet transform or not.

Image Preprocessing. First using the illumination compensation on the original face images (Fig.5) in the database, namely doing gray adjustment and normalization, the images (Fig.6) after transform are obviously clearer than the former ones and helpful for analysis and recognition. Put the one layer Red-Black wavelet transform on the compensation images, then withdraw the images of the blue block sub-band which are the low-dimension approximating images of the original images (Fig.7). Appling Red-Black wavelet transform on the images plays an important role in obscuring the impacts of face expressions and postures and achieved good effects of reducing dimensions.

Table 2. Face recognition precision and time on different models

Methods	PCA+Red-Black wavelet transform	PCA+ illumination compensation	PCA+ illumination compensation+ Red-Black wavelet transform
Recognition rates	75%	85%	93%
Training time(s)	10	23	10
Recognition time(s)	0.28	0.21	0.28

Feature Extraction and Mathing Results. After the image preprocessing, we adopt PCA to extract features and recognize. This paper analyses the results of three models which separately are PCA combined with Red-Black wavelet transform, illumination compensation, Red-Black wavelet transform and illumination compensation. The recognition rates, training and recognition time of different models are showed in the Table 2. We can see that withdrawing the blue block sub-band can obviously reduces the dimensions of the image vector and computation. The reducing training time shows that the low resolution subgraph reduces the computational complexity of the PCA through the Red-Black wavelet transform. Since illumination have great influence on feature extraction and recognition based on PCA, so the recognition effects can be enhanced by illumination compensation. Therefore, the combination of Red-Black wavelet transform, illumination compensation and PCA can achieve more satisfactory

system performance. Comparing with the traditional method using wavelet transform and PCA , the recognition rate is enhanced obviously.

5 Conclusion

The Red-Black wavelet transform divides the rectangular grid digital image into red and black blocks and uses the two-dimensional lifting form to construct sub-band. It is an effective way to wipe off the image relativity and gets the more sparser image. PCA extracts eigenvector on the basis of the whole face grayscale relativity. The eigenvector can retain the main classified information in the original image space and rebuild the original image at the lowest MSE. The method of extracting face features through combining the Red-Black wavelet transform, illumination compensation and PCA can effectively reduce computational complexity at the guarantee of certain recognition rate. Experimental results prove that this scheme is effective in the face recognition.

Acknowledgement. This project is supported by National Science Foundation of China (No. 60572058) and Excellent Young Scholars Research Fund of Beijing Institute of Technology (No. 2006Y0104).

References

1. Chellappa, R., Wilson, C., Sirohe, S.: Human and Machine Recognition of Faces: A Survey. Proceedings of the IEEE 83(5), 705–740 (1995)
2. Yao, R., Huang, J., Wu, X.Q.: An Improved Image Enhancement Algorithm of Histogram Equalization. Journal of The China Railway Society 19(6), 78–81 (1997)
3. Uytterhoeven, G., Bultheel, A.: The Red-Black Wavelet Transform. Katholieke Universiteit Leuven, Belgium (1977)
4. Uytterhoeven, G., Bultheel, A.: The Red-Black Wavelet Transform and the Lifting Scheme. Katholieke Universiteit Leuven, Belgium (2000)
5. Sweldens, W.: The lifting scheme: a Custom-Design Construction of Biorthogonal Wavelets. Applied and Computational Harmonic Analysis 3(2), 186–200 (1996)
6. Kovacvic, J., Sweldens, W.: Wavelet Families of Increasing order in Arbitrary Dimension. IEEE Transactions on Image Processing 9(3), 280–496 (2000)
7. Turk, M., Pentland, A.: Eigenfaces for Recognition. Cognitive Neuroscience 3(1), 71–86 (1991)
8. Stan, Z., Li, J., Lu: Face Recognition Using the Nearest Feeature Line Method. IEEE Trans. on Neural Network 10(2), 439–443 (1999)
9. Wang, H.: Improvement on Red-Black Wavelet Transform. Journal of zaozhuang university 24(5), 61–63 (2007)

Automatic Straight Line Detection through Fixed-Point BYY Harmony Learning

Jinwen Ma* and Lei Li

Department of Information Science, School of Mathematical Sciences and LAMA, Peking University, Beijing, 100871, China
jwma@math.pku.edu.cn

Abstract. Straight line detection in a binary image is a basic but difficult task in image processing and machine vision. Recently, a fast fixed-point BYY harmony learning algorithm has been established to efficiently make model selection automatically during the parameter learning on Gaussian mixture. In this paper, we apply the fixed-point BYY harmony learning algorithm to learning the Gaussians in the dataset of a binary image and utilize the major principal components of the covariance matrices of the estimated Gaussians to represent the straight lines in the image. It is demonstrated well by the experiments that this fixed-point BYY harmony learning approach can both determine the number of straight lines and locate these straight lines accurately in a binary image.

Keywords: Straight line detection, Bayesian Ying-Yang (BYY) harmony learning, Gaussian mixture, Automated model selection, Major principal component.

1 Introduction

Detecting straight lines from a binary image is a basic task in image processing and machine vision. In the pattern recognition literature, a variety of algorithms have been proposed to solve this problem. The Hough transform (HT) and its variations (see Refs. [1,2] for reviews) might be the most classical ones. However, this kind of algorithms usually suffer from large time and space requirements, and detection of false positives, even if the Random Hough Transform (RHT) [3] and the constrained Hough Transform [4] have been proposed to overcome these weaknesses. Later on, there appeared many other algorithms for straight line or curve detection (e.g., [5,6]), but most of these algorithms need to know the number of straight lines or curves in the image in advance.

With the development of the Bayesian Ying-Yang (BYY) harmony learning system and theory [7,8,9,10], a new kind of learning algorithms [11,12,13,14,15] have been established for the Gaussian mixture modeling with a favorite feature that model selection can be made automatically during parameter learning. From the view of line detection, a straight line can be recognized as the major

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 569–576, 2008.
© Springer-Verlag Berlin Heidelberg 2008

principal component of the covariance matrix of certain flat Gaussian of black pixels since the number of black pixels along it is always limited in the image. In such a way, these BYY harmony learning algorithms can learn the Gaussians from the image data automatically and detect the straight lines with the major principal components of their covariance matrices. On the other hand, from the BYY harmony learning on the mixture of experts in [16], a gradient learning algorithm was already proposed for the straight line or ellipse detection, but it was applicable only for some simple cases.

In this paper, we apply the fixed-point BYY harmony learning algorithm [15] to learning an appropriate number of Gaussians and utilize the major principal components of the covariance matrices of these Gaussians to represent the straight lines in the image. It is demonstrated well by the experiments that this fixed-point BYY harmony learning approach can efficiently determine the number of straight lines and locate these straight lines accurately in a binary image.

In the sequel, we introduce the fixed-point BYY harmony learning algorithm and present our new straight line detection approach in Section 2. In Section 3, several experiments on both the simulation and real images are conducted to demonstrate the efficiency of our BYY harmony learning approach. Finally, we conclude briefly in Section 4.

2 Fixed-Point BYY Harmony Learning Approach for Automatic Straight Line Detection

2.1 Fixed-Point BYY Harmony Learning Algorithm

As a powerful statistical model, Gaussian mixture has been widely applied in the fields of information processing and data analysis. Mathematically, the probability density function (pdf) of the Gaussian mixture model of k components in $\Re^d$ is given as follows:

$$\Phi(x) = \sum_{i=1}^{k} \alpha_i q(x|\theta_i), \qquad \forall x \in \Re^d, \tag{1}$$

where $q(x|\theta_i)$ is a Gaussian pdf with the parameters $\theta_i = (m_i, \Sigma_i)$, being given by

$$q(x|\theta_i) = q(x|m_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-m_i)^T \Sigma_i^{-1}(x-m_i)}, \tag{2}$$

and $\alpha_i (\geq 0)$ are the mixing proportions under the constraint $\sum_{i=1}^{k} \alpha_i = 1$. If we encapsulate all the parameters into one vector: $\Theta_k = (\alpha_1, \alpha_2, \ldots, \alpha_k, \theta_1, \theta_2, \ldots, \theta_k)$, then, according to Eq.(1), the pdf of the Gaussian mixture can be rewritten as:

$$\Phi(x|\Theta_k) = \sum_{i=1}^{k} \alpha_i q(x|\theta_i) = \sum_{i=1}^{k} \pi_i q(x|m_i, \Sigma_i). \tag{3}$$

For the Gaussian mixture modeling, there have been several statistical learning algorithms, including the EM algorithm [17] and the k-means algorithm [18]. However, these approaches require an assumption that the number of Gaussians in the mixture is known in advance. Unfortunately, this assumption is practically unrealistic for many unsupervised learning tasks such as clustering or competitive learning. In such a situation, the selection of an appropriate number of Gaussians must be made jointly with the estimation of the parameters, which becomes a rather difficult task [19].

In fact, this model selection problem has been investigated by many researchers from different aspects. The traditional approach was to choose a best number k^* of Gaussians in the mixture via certain model selection criterion, such as Akaike's information criterion (AIC) [20] and the Bayesian Information Criterion (BIC) [21]. However, all the existing theoretic selection criteria have their limitations and often lead to a wrong result. Moreover, the process of evaluating a information criterion or validity index incurs a large computational cost since we need to repeat the entire parameter estimation process at a large number of different values of k. In the middle of 1990s, there appeared some stochastic approaches to infer the mixture model. The two typical approaches are the methods of reversible jump Markov chain Monte Carlo (RJMCMC) [22] and the Dirichlet processes [23], respectively. But these stochastic simulation methods generally require a large number of samples with different sampling methods, not just a set of sample data. Actually, it can efficiently solved through the BYY harmony learning on a BI-architecture of the BYY learning system related to the Gaussian mixture. Given a sample data set $\mathcal{S} = \{x_t\}_{t=1}^N$ from a mixture of k^* Gaussians, the BYY harmony learning for the Gaussian mixture modeling can be implemented by maximizing the following harmony function:

$$J(\Theta_k) = \frac{1}{N} \sum_{t=1}^{N} \sum_{j=1}^{k} \frac{\alpha_j q(x_t \mid \theta_j)}{\sum_{i=1}^{k} \alpha_i q(x_t \mid \theta_j)} ln[\alpha_j q(x_t \mid \theta_j)] \tag{4}$$

where $q(x \mid \theta_j)$ is a Gaussian mixture density given by Eq.(2).

For implementing the maximization of the harmony function, some gradient learning algorithms as well as an annealing learning algorithm were already established in [11,12,13,14]. More recently, a fast fixed-point learning algorithm was proposed in [15]. It was demonstrated well by the simulation experiments on these BYY harmony learning algorithms that as long as k is set to be larger than the true number of Gaussians in the sample data, the number of Gaussians can be automatically selected for the sample data set, with the mixing proportions of the extra Gaussians attenuating to zero. That is, these algorithms owns a favorite feature of automatic model selection during the parameter learning, which was already analyzed and proved for certain cases in [24]. For automatic straight line detection, we will apply the fixed-point BYY harmony learning algorithm to maximizing the harmony function via the following iterative procedure:

$$\alpha_j^+ = \frac{\sum_{t=1}^N h_j(t)}{\sum_{i=1}^i \sum_{t=1}^N h_i(t)}; \tag{5}$$

$$m_j^+ = \frac{1}{\sum_{t=1}^N h_j(t)} \sum_{t=1}^N h_j(t) x_t; \tag{6}$$

$$\Sigma_j^+ = \frac{1}{\sum_{t=1}^N h_j(t)} \sum_{t=1}^N h_j(t)(x_t - \hat{m_j})(x_t - \hat{m_j})^T, \tag{7}$$

where $h_j(t) = p(j|x_t) + \sum_{i=1}^k p(j|x_t)(\delta_{ij} - p(j|x_t))ln[\alpha_i q(x_t|m_i, \Sigma_i)]$, $p(j|x_t) = \alpha_j q(x_t|m_j, \Sigma_j)/\sum_{i=1}^k \alpha_i q(x_t|m_i, \Sigma_i)$ and δ_{ij} is the Kronecker function. It can be seen from Eqs (5)-(7) that the fixed-point BYY harmony learning algorithm is very similar to the EM algorithm for Gaussian mixture. However, since $h_j(t)$ introduces a rewarding and penalizing mechanism on the mixing proportions [13], it differs from the EM algorithm and owns the favorite feature of automated model selection on Gaussian mixture.

2.2 The Proposed Approach to Automatic Straight Line Detection

Given a set of black points or pixels $\mathcal{B} = \{x_t\}_{t=1}^N$ ($x_t = [x_{1,t}, x_{2,t}]^T$) in a binary image, we regard the black points along each line as one flat Gaussian distribution. That is, those black points can be assumed to be subject to a 2-dimensional Gaussian mixture distribution. Then, we can utilize the fixed-point BYY harmony learning algorithm to estimate those flat Gaussians and use the major principal components of their covariance matrices to represent the straight lines as long as k is set to be larger than the number k^* of the straight lines in the image. In order to speed up the convergence of the algorithm, we set a threshold value $\delta > 0$ such that as soon as the mixing proportion is lower than δ, the corresponding Gaussian will be discarded from the mixture.

With the convergence of the fixed-point BYY harmony learning algorithm on $\mathcal{B}$ with $k \geq k^*$, we get k^* flat Gaussians with the parameters $\{(\alpha_i, m_i, \Sigma_i)\}_{i=1}^{k^*}$ from the resulted mixture. Then, from each Gaussian $(\alpha_i, m_i, \Sigma_i)$, we pick up m_i and the major principle component $V_{1,i}$ of Σ_i to construct a straight line equation $l_i : U_{1,i}^T(x - m_i) = 0$, where $U_{1,i}$ is the unit vector being orthogonal to $V_{1,i}$, with the mixing proportion α_i representing the proportion of the number of points along this straight line l_i. Since the sample points are in a 2-dimensional space, $U_{1,i}$ can be uniquely determined and easily solved from $V_{1,i}$, without considering its direction. Hence, the problem of detecting multiple straight lines in a binary image has been turned into the Gaussian mixing modeling problem of both model selection and parameter learning, which can be efficiently solved by the fixed-point BYY harmony learning algorithm automatically.

With the above preparations, as $k(> k^*)$, the stop criterion threshold value $\varepsilon(> 0)$ and the component annihilation threshold value $\delta(> 0)$ are all prefixed, the procedure of our fixed-point BYY harmony learning approach to automatic straight line detection with $\mathcal{B}$ can be summarized as follows.

1. Let $t = 1$ and set the initial parameters Θ_0 of the Gaussian mixture as randomly as possible.
2. At time t, update the parameters of the Gaussian mixture at time $t-1$ by Eqs (5)-(7) to get the new parameters $\Theta_t = (\alpha_i, m_i, \Sigma_i)_{i=1}^k$;
3. If $|J(\Theta_t) - J(\Theta_{t-1})| \leq \varepsilon$, stop and get the result Θ_t, and go to Step 5; otherwise, let $t = t + 1$ and go to Step 4.
4. If some $\alpha_i \leq \delta$, discard the component $\theta_i = (\alpha_i, m_i, \Sigma_i)$ from the mixture and modify the mixing proportions with the constraint $\sum_{j=1}^k \alpha_j = 1$. Return to Step 2.
5. Pick up m_i and the major principle component $V_{1,i}$ of Σ_i of each Gaussian $(\alpha_j, m_j, \Sigma_j)$ in the resulted mixture to construct a straight line equation $l_i : U_{1,i}^T(x - m_i) = 0$.

It can be easily found from the above automatic straight line detection procedure that the operation of the fixed-point BYY harmony learning algorithm tries to increase the total harmony function on the Gaussian mixture so that the extra Gaussians or corresponding straight lines will be discarded automatically during the parameter learning or estimation.

3 Experimental Results

In this section, several simulation and practical experiments are conducted to demonstrate the efficiency of our proposed fixed-point BYY harmony learning approach. In all the experiments, the initial means of the Gaussians in the mixture are trained by the k-means algorithm on the sample data set $\mathcal{B}$. Moreover, the stop criterion threshold value ε is set to be $10 * e^{-8}$ and the component annihilation threshold value δ is set to be 0.08. For clarity, the original and detected straight lines will be drawn with red color, but the sample points along different straight lines will be drawn in black.

3.1 Simulation Results

For testing the proposed approach, simulation experiments are conducted on three binary image datasets consisting of different numbers of straight lines, which are shown in Fig.1(a),(b),(c), respectively. We implement the fixed-point BYY harmony learning algorithm on each of these datasets with $k = 8$. The results of the automatic straight line detection on the three image datasets are shown in Fig.1(d),(e),(f), respectively. Actually, in each case, some random noise from a Gaussian distribution with zero mean and a standard variance 0.2 is added to the coordinates of each black point. It can be seen from the experimental results that the correct number of straight lines is determined automatically to match the actual straight lines accurately in each image dataset.

3.2 Automatic Container Recognition

Automatic container recognition system is very useful for customs or logistic management. In fact, our proposed fixed-point BYY harmony learning approach

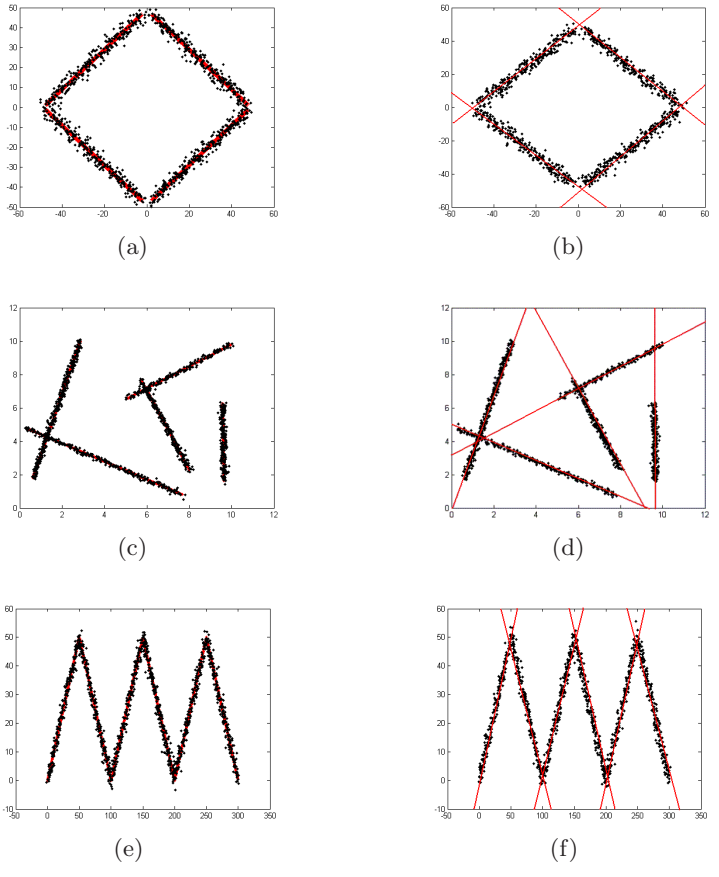


Fig. 1. The experiments results of the automatic straight line detection through the fixed-point BYY harmony learning approach. (a),(b),(c) are the three binary image datasets, while (d),(e),(f) are their results of the straight line detection.

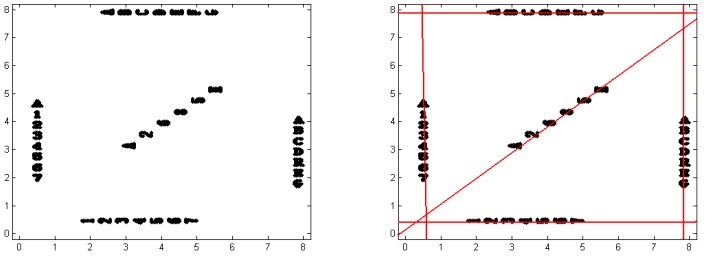


Fig. 2. The experiments results on automatic container recognition. (a) The original container image with five series of numbers (with letters). (b) The result of the automatic container recognition through the fixed-point BYY harmony learning approach.

can be applied to assisting to establish such a system. Container recognition is usually based on the captured container number located at the back of the container. Specifically, the container, as shown in Fig.2(a), can be recognized by the five series of numbers (with letters). The recognition process consists of two steps. The first step is to locate and extract each rectangular area in the raw image that contains a series of numbers, while the second step is to actually recognize these numbers via some image processing and pattern recognition techniques.

For the first step, we implement the fixed-point BYY learning algorithm to roughly locate the container numbers via detecting the five straight lines through the five series of the numbers, respectively. As shown in Fig.2(b), these five straight lines can locate the series of numbers very well. Based on the detected strip lines, we can extract the rectangular areas of the numbers from the raw image. Finally, the numbers can be subsequently recognized via some image processing and pattern recognition techniques.

4 Conclusions

We have investigated the straight line detection in a binary image from the point of view of the Bayesian Ying-Yang (BYY) harmony learning and proposed the fixed-point BYY harmony learning approach to automatic straight line detection. Actually, we implement the fixed-point BYY harmony learning algorithm to learn a number of flat Gaussians from an image dataset automatically to represent the black points along the actual straight lines, respectively, and locate these straight lines with the major principal components of the covariance matrices of the obtained Gaussians. It is demonstrated well by the experiments on the simulated and real-world images that this fixed-point BYY harmony learning approach can both determine the number of straight lines and locate these straight lines accurately in an image.

Acknowledgments. This work was supported by the Natural Science Foundation of China for grant 60771061.

References

1. Ballard, D.: Generalizing the Hough Transform to Detect Arbitrary Shapes. Pattern Recognition 13(2), 111–122 (1981)
2. Illingworth, J., Kittler, J.: A Survey of the Hough Transform. Computer Vision, Graphs, & Image Process 44, 87–116 (1988)
3. Xu, L., Oja, E., Kultanen, P.: A New Curve Detection Method: Randomized Hough Transform (RHT). Pattern Recognition Letter 11, 331–338 (1990)
4. Olson, C.F.: Constrained Hough Transform for Curve Detection. Computer Vision and Image Understanding 73(3), 329–345 (1999)
5. Olson, C.F.: Locating Geometric Primitives by Pruning the Parameter Space. Pattern Recognition 34(6), 1247–1256 (2001)
6. Liu, Z.Y., Qiong, H., Xu, L.: Multisets Mixture Learning-based Ellipse Detection. Pattern Recognition 39, 731–735 (2006)

7. Xu, L.: Ying-Yang machine: A Bayesian-Kullback Scheme for Unified Learnings and New Results on Vector Quantization. In: Proc. 1995 Int. Conf. on Neural Information Processing (ICONIP 1995), vol. 2, pp. 977–988 (1995)
8. Xu, L.: Bayesian Ying-Yang Machine, Clustering and Number of Clusters. Pattern Recognition Letters 18, 1167–1178 (1997)
9. Xu, L.: Best Harmony, Unified RPCL and Automated Model Selection for Unsupervised and Supervised Learning on Gaussian Mixtures, Three-layer Nets and ME-RBF-SVM Models. International Journal of Neural Systems 11(1), 43–69 (2001)
10. Xu, L.: BYY Harmony Learning, Structural RPCL, and Topological Self-organzing on Mixture Modes. Neural Networks 15, 1231–1237 (2002)
11. Ma, J., Wang, T., Xu, L.: A Gradient BYY Harmony Learning Rule on Gaussian Mixture with Automated Model Selection. Neurocomputing 56, 481–487 (2004)
12. Ma, J., Gao, B., Wang, Y., Cheng, Q.: Conjugate and Natural Gradient Rules for BYY Harmony Learning on Gaussian Mixture with Automated Model Selection. International Journal of Pattern Recognition and Artificial Intelligence 19, 701–713 (2005)
13. Ma, J., Wang, L.: BYY Harmony Learning on Finite Mixture: Adaptive Gradient Implementation and A Floating RPCL Mechanism. Neural Processing Letters 24(1), 19–40 (2006)
14. Ma, J., Liu, J.: The BYY annealing learning algorithm for Gaussian mixture with automated model selection. Pattern Recognition 40, 2029–2037 (2007)
15. Ma, J., He, X.: A Fast Fixed-point BYY Harmony Learning Algorithm on Gaussian Mixture with Automated Model Selection. Pattern Recognition Letters 29(6), 701–711 (2008)
16. Lu, Z., Cheng, Q., Ma, J.: A Gradient BYY Harmony Learning Algorithm on Mixture of Experts for Curve Detection. In: Gallagher, M., Hogan, J.P., Maire, F. (eds.) IDEAL 2005. LNCS, vol. 3578, pp. 250–257. Springer, Heidelberg (2005)
17. Render, R.A., Walker, H.F.: Mixture Densities, Maximum Likelihood and the EM Algorithm. SIAM Review 26(2), 195–239 (1984)
18. Jain, A.K., Dubes, R.C.: Algorithm for Clustering Data. Prentice Hall, Englewood Cliffs (1988)
19. Hartigan, J.A.: Distribution Problems in Clustering. In: Van Ryzin, J. (ed.) Classification and Clustering, pp. 45–72. Academic Press, New York (1977)
20. Akaike, H.: A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control AC-19, 716–723 (1974)
21. Scharz, G.: Estimating the Dimension of a Model. The Annals of Statistics 6, 461–464 (1978)
22. Green, P.J.: Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. Biometrika 82(4), 711–732 (1995)
23. Escobar, M.D., West, M.: Bayesian Density Estimation and Inference Using Mixtures. Journal of the American Statistical Association 90(430), 577–588 (1995)
24. Ma, J.: Automated Model Selection (AMS) on Finite Mixtures: A Theoretical Analysis. In: Proc. 2006 International Joint Conference on Neural Networks (IJCNN 2006), Vancouver, Canada, July 16-21, 2006, pp. 8255–8261 (2006)

A Novel Solution for Surveillance System Based on Skillful Image Sensor

Tianding Chen

Institute of Communications and Information Technology, Zhejiang Gongshang University, Hangzhou 310018
chentianding@163.com

Abstract. Omni-camera is almost near or even beyond a full hemisphere, it is popularly applied in the fields of visual surveillance, and vision-based robot or autonomous vehicle navigation. The captured omni-directional image should be rectified into a normal perspective-view or panoramic image for convenient human viewing or image-proof preservation. It develops a real-time system to supervise the situation of the indoor environments. This surveillance system combines a panoramic video camera and three active PTZ cameras. We use the omnidirectional video that is caught by panoramic video camera as input to track the moving subjects in an indoor room. We can automatically track the moving subjects or manually find the moving subjects. The coordinate of each subject in the indoor room is estimated in order to direct the attention of one active pan-tilt-zoom camera. When we find the moving subjects automatically or manually, this system can automatically control these PTZ cameras to obtain the high-resolution images and video sequences of moving subjects. Compared with the other kinds of tracking systems, our system can more easily enlarge the monitoring area by adding cameras. The experimental results show that those moving object can be tracked exactly and can be integrated with the guiding system very well.

Keywords: Omni-camera, Surveillance system, Moving subject.

1 Introduction

In the modern society of scientific century, there are many surveillance systems to supervise the situation of the indoor environment such as bookstore, bank, convenience store, saloon, and airport. There are two common types of surveillance systems: (1) systems without human monitoring and merely recording and storing video footage, and (2) systems requiring continuous human monitoring. The two kinds of surveillance systems have some drawbacks. For example, when a burglar invades, the surveillance system without human monitoring can not issue the alarms in real-time. Whereas the systems with continuous human monitoring need human to monitor incessantly, so this system takes much manpower to issue the alarms in real-time. In order to overcome these flaws, a real-time system which can automatically analyze the serial digital videos to detect the moving subjects in indoor room is required.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 577–584, 2008.
© Springer-Verlag Berlin Heidelberg 2008

It is a kernel concept in computer vision to identify an object in a working space by analyzing an image captured by a camera. The technique of building the relationship between the coordinate system of a working space and that of a captured image is called camera calibration. The work of transforming a pixel in a captured image into a corresponding point in a working space is called image transformation. The image transformation work relies on the completion of camera calibration by using the calibrated parameters created by the calibration procedure.

For a traditional perspective camera, calibration is a well-known and mature technique [1]. But for a kind of so-called omni-directional camera, the camera calibration work is a new and developing technique, and many practical problems need to be resolved. Even in the field of traditional perspective camera calibration, there are still some problems in real applications. In this study, we investigate the image unwarping problem for omni-directional cameras and propose some new solutions to such a kind of problem. To conduct image unwarping which is an image transformation problem, we have to conduct the camera calibration work in advance. For some real applications using perspective cameras, for example, applying an image-based pointing device, we will propose some new ideas for improving the accuracy of the calibration result.

It is well known in computer vision that enlarging the field of view (FOV) of a camera enhances the visual coverage, reduces the blind area, and saves the computation time, of the camera system, especially in applications like visual surveillance and vision-based robot or autonomous vehicle navigation. An extreme way is to expand the FOV to be beyond a full hemisphere by the use of some specially designed optics. A popular name for this kind of camera is omni-directional camera [2], and an image taken by it is called an omni-directional image. Omni-cameras can be categorized into two types according to the involved optics, namely, dioptric and catadioptric. A dioptric omni-camera captures incoming light going directly into the imaging sensor to form omni-images. Examples of image unwarping works for a kind of dioptric omni-camera called fish-eye camera can be found in [3][4]. A catadioptric omni-camera [5] captures incoming light reflected by a mirror to form omni-images.

2 Vision System Setup and Camera Calibration

The proposed vision system for intelligent surveillance is mounted on the ceiling of an indoor environment. It contains a CCD camera with a PAL (Panoramic Annular Lens) optical system (see Fig. and 1) [6], a frame grabber, and a PC. The video sequences are transferred to the PC at a frame rate of 12 fps with image resolution of 320 × 240 pixels.

The PAL optics consists of a single glass block with two reflective and two refractive planes, and has a single center of projection. Thus, the captured concentric omnidirectional image can be unwarped to create a panoramic image provided that the equivalent focal length is available. Camera calibration of the omnidirectional vision system includes two parts– one is to derive the equivalent focal length and image center, and the other one is to find the one-to-one correspondence between the ground positions and the image pixels. Although the geometry of the PAL imaging system is complex, it can be modeled by the cylindrical coordinate system with a

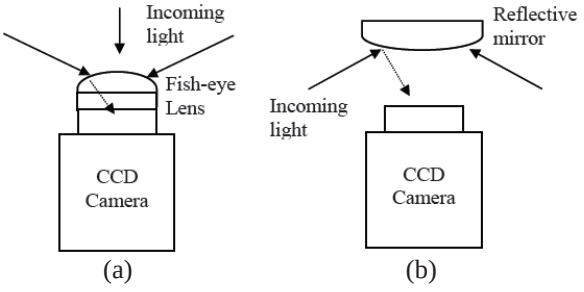


Fig. 1. Structures of omni-cameras. (a) Dioptric. (b) Catadioptric.

single virtual viewpoint located at the original of the coordinate system. The panoramic image corresponding to the "virtual camera" is given by the cylinder with radius of the effective focal length. The center of the annular image, i.e., the viewpoint or origin of the virtual camera, is determined by the intersection of vertical edges in the scene. In the implementation, several upright poles were hung from the ceiling, the Sobel detector was used to obtain an edge image, followed by Hough transform for robust line fitting. To increase the stability, a sequence of images was recorded for camera calibration and the mode of the distribution from the images was given as the image center. Fig.2 shows the image center obtained using three poles hung from the ceiling.

Fig. 2. The calibration result of hypercatadioptric camera used in this study

It is given by the intersection of the three edges in the annular image. To obtain the effective focal length of the PAL imaging system, upright calibration objects (poles) with known metric are placed at several fixed locations. Suppose the height of a pole is h and its location is d from the ground projection of the camera, then the effective focal length is given by

$$f = \frac{dvs}{h} \tag{1}$$

where v is the size of the object in pixel, and s is the pixel size of the CCD in the radial direction. Due to the characteristics of the PAL sensors, the effective focal length obtained from Eq. (1) is usually not a constant for different object sizes and

locations. In the experiments, the calibration pattern was placed upright such that the projection of its midpoint appeared in the central row of the panoramic image. However, it should be noted that the image scanline does not necessarily correspond to the middle concentric circle in the PAL image. The second part of omnidirectional camera calibration is to establish the relationship between the image pixels and the 3-D ground positions. Generally speaking, 3-D to 2-D projection is an ill-posed problem and cannot be recovered from a single viewpoint. However, if the object possesses additional geometric constraints such as coplanarity, collinearity, it is possible to create a lookup table with the 3-D coordinate of each image pixel. Since the lookup tables are not identical for different omnidirectional camera position and orientation, they should be obtained for the individual environments prior to any surveillance tasks.

3 Solution of Omnidirectional Visual Tracking

Omni-cameras are used in many applications such as visual surveillance and robot vision [7][8] for taking omni-images of camera surroundings. Usually, there exists an extra need to create perspective-view images from omni-images for human comprehensibility or event recording. This image unwarping work usually is a complicated work.

Omni-cameras can be categorized into two types according to the involved optics, namely, dioptric and catadioptric, as mentioned previously [5]. A dioptric omni-camera captures incoming light going directly into the imaging sensor to form omni-images. An example of dioptric omni-cameras is the fish-eye camera. A catadioptric omni-camera captures incoming light reflected by a mirror to form omni-images. The mirror surface may be in various shapes, like conic, hyperbolic, etc.

The method for unwarping omni-images is different for each distinct type of omni-camera. Generally speaking, omni-images taken by the SVP catadioptric camera [5] as well as the dioptric camera are easier to unwarp than those taken by the non-SVP catadioptric camera [9]. Conventional methods for unwarping omni-images require the knowledge of certain camera parameters, like the focal length of the lens, the coefficients of the mirror surface shape equation, etc to do calibration before the unwarping can be done [10][11]. In the last chapter, we have proposed a method of this kind. But in some situations, we cannot get the complete information of the omni-camera parameters. Then the unwarping work cannot be conducted. It is desired to have a more convenient way to deal with this problem. The proposed method consists of three major stages: (1) landmark learning, (2) table creation, and (3) image unwarping.

A. Landmark learning

The first step, landmark learning, is a procedure in which some pairs of selected world space points with known positions and their corresponding pixels in a taken omni-image are set up. More specifically, the coordinates of at least five points, called landmark points hereafter, which are easy to identify in the world space (for example, a corner in a room), are measured manually with respect to a selected origin in the world space. Then the corresponding pixels of such landmark points in the taken omni-image are segmented out. A world space point and its corresponding image pixel so selected together are said to form a landmark point pair.

B. Table creation

The second step, table creation, is a procedure in which a pano-mapping table is built using the coordinate data of the landmark point pairs. The table is 2-dimensional in nature with the horizontal and vertical axes specifying respectively the range of the azimuth angle θ as well as that of the elevation angle ρ of all possible incident light rays going through the mirror center. An illustration is shown in Fig. 3.

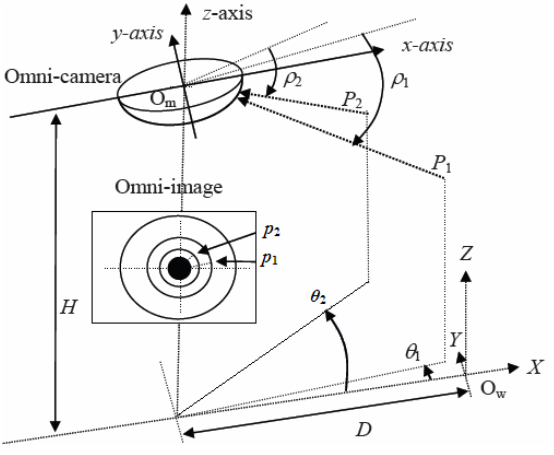


Fig. 3. System configuration

Each entry E_{ij} with indices (i, j) in the pano-mapping table specifies an azimuth-elevation angle pair (θ_i, ρ_j), which represent an infinite set S_{ij} of world space points passing through by the light ray with azimuth angle θ_i and elevation angle ρ_j. These world space points in S_{ij} are all projected onto an identical pixel p_{ij} in any omni-image taken by the camera, forming a pano-mapping f_{pm} from S_{ij} to p_{ij}. This mapping is shown in the table by filling entry E_{ij} with the coordinates (u_{ij}, v_{ij}) of pixel p_{ij} in the omni-image. The table as a whole specifies the nature of the omni-camera, and may be used to create any panoramic or perspective-view images, as described subsequently.

C. Image unwarping

The third step, image unwarping, is a procedure in which the pano-mapping table T_{pm} of an omni-camera is used as a media to construct a panoramic or perspective-view image Q of any size for any viewpoint from a given omni-image I taken by the omni-camera. The basic concept in the procedure is to map each pixel q in Q to an entry E_{ij} with coordinate values (u_{ij}, v_{ij}) in the pano-mapping table T_{pm} and to assign the color value of the pixel at coordinates (u_{ij}, v_{ij}) of image I to pixel q in Q.

A new approach to unwarping of omni-images taken by all kinds of omni-cameras is proposed. The approach is based on the use of pano-mapping tables proposed in this study, which may be regarded as a summary of the information conveyed by all the parameters of an omni-camera. The pano-mapping table is created once forever for each omni-camera, and can be used to construct panoramic or perspective-view

images of any viewpoint in the world space. This is a simple and efficient way to unwarping an omni-image taken by any omni-camera, when some of the physical parameters of the omni-camera cannot be obtained. One of table creation procedure algorithm is filling entries of pano-mapping table. The result is shown in Fig. 4.

Algorithm. Filling entries of pano-mapping table.

Step1. Divide the range 2π of the azimuth angles into M intervals, and compute the i-th azimuth angle θ_i by

$$\theta_i = i \times (2\pi / M), \text{ for } \quad i = 0,1,\ldots,M-1 \tag{2}$$

Step 2. Divide the range $[\rho_e - \rho_z]$ of the elevation angles into N intervals, and compute the j-th elevation angle ρ_j by

$$\rho_j = j \times [(\rho_e - \rho_z)/N] + \rho_z, \text{ for } \quad j = 0,1,\ldots,N-1 \tag{3}$$

Step 3. Fill the entry E_{ij} with the corresponding image coordinates (u_{ij}, v_{ij}) computed according to Eqs. (4) and (5) as follows:

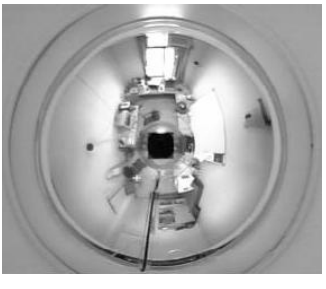
$$u_{ij} = r_j \times \cos\theta_i \tag{4}$$

$$v_{ij} = r_j \times \sin\theta_i \tag{5}$$

where r_j is computed by

$$r_j = f_r(\rho_j) = a_0 + a_1 \times \rho_j^1 + a_2 \times \rho_j^2 + a_3 \times \rho_j^3 + a_4 \times \rho_j^4 \tag{6}$$

with $(a_0, a_1, a_2, a_3, a_4)$ being those computed by estimation of coefficients of radial stretching function algorithm.



(a) image sensor camera

(b) result of tracking scene

Fig. 4. System result

4 Behavior Analysis

In the home or office environments, there are typically several places where people spend most of their time without significant motion changes, such as chairs, beds, or the regions close to the television or computers. These places can be marked as “inactivity zones” and usually have little changes in the video surveillance. On the other hand, there are some places involve frequent changes in the scene such as the hallway, entrance to the room, and free space in the environment. These regions are

referred to as “activity zones” since most of the detectable motion changes happen in those areas. There is also a third case called “non-activity zones”, which indicate that any normal human activities should not occur in the regions. Based on the above classification, the surveillance environment is partitioned into three regions. An abnormal behavior warning system is established according to the activities happened in the different zones. It is clear that the size of an object appeared in the image depends on the location of the object, even for the omnidirectional imaging sensors. Based on this fact and the above calibration information, the projected height of an object (or a person) with known physical dimension can be computed for different ground position in a concentric manner. Thus, fall detection of the target can be given by identifying its standing position (the bottom) and its top. By comparing the object length appeared in the image and the computed length according to the object’s physical height, some activities of the object such as falling, sitting, etc. can be detected. To make the fall detection algorithm more robust and to avoid false alarm, the motion vectors obtained from optical flow are used to detect the sudden changes of the object’s height and motion direction.

5 Conclusions

Intelligent surveillance is one of the important topics in the nursing and home-care robotic systems. This paper presented a system which tracks all subjects in the indoor room using panoramic video camera, and uses this information to guide the visual attention of active PTZ camera toward objects of interest, and proposed a video surveillance system using an omnidirectional CCD camera. Automatic object detection, people tracking, and fall detection have been implemented. In the future, we have a method of stereo matching of planar scenes will be proposed. First, the homography of each is obtained and then we use homograpy to segment each plane. After that, two images have to be rectified so that the corresponding search would change from 2D to 1D. So, a PTZ camera will be combined with the omnidirectional tracking system to obtain a closer look of the target for recognition purposes. The active PTZ cameras provide high quality video which could be used for face and action recognition.

Acknowledgements. This research work is supported by Zhejiang Province Nature Science Foundation under Grant Y107411.

References

1. Salvi, J., Armangué, X., Batle, J.: A Comparative Review of Camera Calibrating Methods with Accuracy Evaluation. Pattern Recognition 35(7), 1617–1635 (2002)
2. Nayar, S.K.: Omnidirectional Vision. In: Proceedings of Eighth International Symposium on Robotics Research (ISRR), Shonan, Japan (1997)
3. Kannala, J., Brandt, S.: A Generic Camera Calibration Method for Fish-Eye Lenses. In: Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, U.K, vol. 1, pp. 10–13 (2004)

4. Shah, S., Aggarwal, J.K.: Intrinsic Parameter Calibration Procedure for A (High-Distortion) Fish-Eye Lens Camera with Distortion Model and Accuracy Estimation. Pattern Recognition 29(11), 1775–1788 (1996)
5. Nayar, S.K.: Catadioptric Omni-directional Camera. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San-Juan, Puerto Rico, pp. 482–488 (1997)
6. Baker, S., Nayar, S.K.: A Theory of Single-Viewpoint Catadioptric Image Formation. International Journal of Computer Vision 35(2), 175–196 (1999)
7. Morita, S., Yamazawa, Yokoya, N.: Networked Video Surveillance Using Multiple Omni-directional Cameras. In: Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation, Kobe, Japan, pp. 1245–1250 (2003)
8. Tang, L., Yuta, S.: Indoor Navigation for Mobile Robots Using Memorized Omni-directional Images and Robot's Motion. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and System, vol. 1, pp. 269–274 (2002)
9. Jeng, S.W., Tsai, W.H.: Precise Image Unwarping of Omnidirectional Cameras with Hyperbolic-Shaped Mirrors. In: Proceedings of 16th IPPR Conference on Computer Vision, Graphics and Image Processing, pp. 414–422 (2003)
10. Strelow, D., Mishler, J., Koes, D., Singh, S.: Precise Omnidirectional Camera Calibration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, vol. 1, pp. 689–694 (2001)
11. Scaramuzza, D., Martinelli, A., Siegwart, R.: A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion. In: Proceedings of the Fourth IEEE International Conference on Computer Vision System (2006)

Collision Prevention for Exploiting Spatial Reuse in Ad Hoc Network Using Directional Antenna

Pu Qingxian and Won-Joo Hwang

Department Information and Communications Engineering
Inje University, 607 Obangdong, Kyungnam, Gimhae, South Korea, 621-749
yunipu@hotmail.com, ichwang@inje.ac.kr

Abstract. To improve the performance of a wireless Ad Hoc network at the MAC layer, one way is to prevent collision and allow spatial reuse among neighbor nodes as much as possible. In this paper, we present a novel MAC protocol that prevents collision and supports spatial reuse. The proposed MAC protocol uses a control window between RTS/CTS packets and DATA/ACK packets which allows neighbor nodes to exchange control packets inserted to realize spatial reuse. Simulation result shows the advantages of our proposal in comparison with existing MAC in terms of network throughput and end-to-end delay.

Keywords: Directional Antenna, MAC, IEEE 802.11.

1 Introduction

Wireless Ad Hoc networks are a self-organizing system of mobile nodes which can be rapidly deployed without any established infrastructure or centralized control facilities. But it is easy to collide with other neighbor nodes. Using directional antenna can exploit spatial reuse and reduce collision. However, existing MAC protocols are not or weakly designed for directional antenna to support spatial reuse.

The previous works on wireless ad hoc networks are primarily designed for a single-hop and omni-directional antennas wireless environment, such as the 802.11 MAC [1]. In such a single-channel [2] environment, the 802.11 MAC contention resolution mechanisms focuses primarily on ensuring that only one node pair receives collision-free access to the channel at any single instant. The 802.11 MAC does not seek to exploit the spatial reuse inherent in multihop networks. By exploiting this spatial reuse, we should be able to significantly increase the number of concurrent transmissions by distinct sender-receiver pairs that are spaced sufficiently apart. An 802.11 MAC protocols in a multi-channel environment, it is easy to occur collisions because of hidden nodes and exposed nodes. Thus, exploiting spatial reuse between neighbor nodes in ad hoc network is the main topic of this paper.

Some researchers in the past have addressed challenges to prevent collision and improve wireless spatial reuse [3]. [4]. [6]. [7]. A. Chandar [1] has presented that IEEE 802.11 DCF uses a 4-way distributed handshake mechanism to resolve contentions between nodes to prevent collision. But it doesn't permit concurrent

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 585–592, 2008.
© Springer-Verlag Berlin Heidelberg 2008

transmission in two nodes which are either neighbors or have a common neighboring node. Conversely, if any node is a receiver, only one node in its one-hop neighborhood is allowed to be a transmitter.

Y. Ko [4] proposed MAC protocol using directional antenna to permit concurrent transmission. In this protocol, transmitter sends a directional RTS and the receiver responses omni-directional CTS. They assume that the transmitter knows the receiver's location, so it transmits directional RTS to it. They propose an alternative of that scheme in case of lack of information for the location of the receiver. In this case the RTS packet is transmitted in omni mode in order to seek the receiver. M. Takai [6] proposed Directional Virtual Carrier Sensing in which they use directional RTS and CTS transmission. They assume that the receiver's location is known by the transmitter. In the opposite situation, they propose the omni-directional transmission of RTS packet. They also propose a cache scheme where they maintain information about the location of their neighbors, which is updated every time a node receives a frame.

In this paper, we propose a novel MAC protocol to handle the issue of exploiting spatial reuse.

- We insert a control window between the RTS/CTS and DATA/ACK. It helps neighbor nodes to exchange control packets to be engaged in concurrent transmission. Based on directional antenna and control window, we combine directional antenna with control window which achieve a high performance.
- We evaluate our protocol performance and compare it with existing MAC protocol in terms of throughput and end-to-end delay.

The rest of this paper is organized as follows. In section 2, we give an antenna model and define the problem. In section 3, we proposed a new MAC protocol to prevent collision and realize spatial reuse. In section 4, we compare the performance of our proposed protocols with IEEE 802.11. And section 6 concludes this paper.

2 Preliminaries

2.1 Antenna Model

In this study, we assume a switched beam antenna system. The assumed antenna model is comprised of M beam patterns. The non-overlapping directional beams are numbered from 1 to M, starting at the three o'clock position and running clockwise. The antenna system offers two modes of operation: Omni and Directional. In Omni mode, a node receives signal from all directional with gain G^o. While the signal is being received, the antenna detects the beam (direction) on which the signal power is strongest and goes into the directional mode. In directional mode, a node can point its beam towards a specific directional with gain G^d.

2.2 Problem Definition

In IEEE 802.11 protocol limits spatial reuse of the wireless channel by silencing all nodes in the neighborhood of the sender and receiver. In Fig. 1, node C and node B are one-hop neighbors, and node A's transmission range does not include node C, and

node D's transmission range does not include node B. For the case (1), it seems that the transmission between node A to node B and the transmission between node C and node D can initiate at the same time, because of node A is not in the node D's transmission range and node A doesn't receive RTS packet from node D. However, node B sends a CTS packet in response to node A's RTS packet, node C is also aware node B's CTS packet. Thus node C cannot respond with a CTS packet to node D. For case (2), node B and node C should be able to transmit to node A and node D respectively at the same time, but IEEE 802.11 does not permit such concurrent transmission, because node B's RTS packet prohibits node C sending RTS packet. For case (3) and case (4), it is clear that concurrent transmission is impossible. Since node B's transmission to node A would collide with the node D's transmission at node C, and node A's transmission to node B would collide with node C's transmission. Therefore, IEEE 802.11 protocol prevents collision but does not support concurrent transmission.

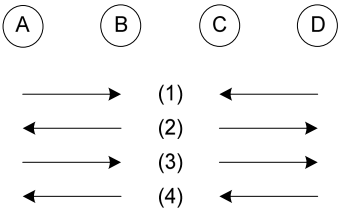


Fig. 1. 802.11 MAC's Failure to Exploit Spatial Reuse

3 Proposed MAC Protocols

3.1 Position Communication Pairs

As we know, ad hoc networks topology is continuously changing due to frequent node movements. Before the communication, we need to determine the position of communication pairs. A MAC protocol refers to common used MAC [7] that uses directional antennas in network where the mobile nodes do not have any location information. The key feature that has been added in the adaptation is a mechanism for the transmitting and receiving nodes to determine the directions of each other. Before the data packet being transmitted, the MAC must be capable of finding the direction of the receiver node. Similarly, as we require the receiver to use directional antennas as well, the receiver node must also know the direction of the transmitter before it receives the transmitted data packet. Any node that wishes to transmitted data packet to a neighbor first sends an omni-directional RTS packet. The receiver receives the RTS packet, and it will estimate the position of the transmitter. After that the receiver sends directional CTS packet at that direction. When transmitter receives CTS packet from receiver, it can estimate the position of receiver.

3.2 Exploit Spatial Reuse

In the previous section we pointed the disadvantages of 802.11 MAC protocols which failures to exploit spatial reuse. To improve the performance of 802.11 MAC

protocols and to solve the problem of case (1) in Fig. 1, the proposed MAC protocol utilized control window realizing the spatial diversity.

Because of a node converting between a transmitter and receiver roles multiple times during a packet transfer without a precision, the 802.11 MAC precludes the possibility of concurrent transmissions by two neighboring nodes that are either both senders or both recipients. In addition, in the 802.11 4-way handshake mechanism, neighbors cannot initialize to be a transmitter while a node pair is transmitting a packet, until the original 4-way handshake is completed. Therefore, only two neighbors being either both transmitters or both receivers and a window which between the RTS/CTS exchange and the subsequent DATA/ACK exchange allows other neighboring pairs to exchange RTS/CTS packets within the control phase window of the first pair, and subsequent pairs to align their DATA/ACK transmission phases with that of the first pair, modified MAC can support concurrent transmissions. The control window is put in place by the first pair (A-to-B). A subsequent RTS/CTS exchange by a neighboring pair (D-to-C) does not redefine the window; subsequent pairs instead use the remaining portion of the control window to align their data transmission with the first pair. Consequently, one-hop neighbors exchange the roles between transmitters and receivers in unison at explicitly defined instants and neighbors synchronize their reception periods. Thus this protocol avoids the problem of packet collisions.

A) One Master Transmission. In Fig. 2 (A), corresponding to case2 in Fig. 1, node B sends a RTS packet to node A, where node B is called the master node and RTS packet sets up the master transmission schedule. If node C overhears this RTS packet and node C has a packet to transmit, it will set up an overlapping transmission to node D, where node C is called the slave node and sets the slave transmission schedule, aligning the starts of the DATA and ACK phases.

B) Two Masters Transmissions. In Fig. 2 (A), node F is neighbor of node C and node B, but node C is not a neighbor of node B. The two transmissions A-to-B and D-to-C have been scheduled, so node F has two masters, node B and node C. When node E sends a RTS packet to node F, the data transmission from node E to node F must align with data transmission from node D to node C, and node F's ACK align with node B's ACK to node A. In general, if a node has more than one master, it has to align the proposed DATA transmission with that of the master with earliest DATA transmission and align the ACK with that of the master with the latest ACK. Thus, all master recipient nodes other than the master with the latest ACK are blocked from scheduling any further receptions till the master transmission with the latest ACK, completes. This means node C cannot schedule any further reception from node D before node F sends its ACK to node E aligned with the ACK from node B to node A. Otherwise, a subsequent CTS packet from node C could interfere with node F's reception of data. Therefore, utilizing control window only is not enough.

3.3 Using Directional CTS

In section 3.1 and section 3.2, we mentioned that utilizing directional antenna model can determine the position of the node which receive or send packets, and control window which exploit spatial reuse. This section details the proposed protocol equipped with directional antenna and control window together. In Fig. 3, node A and

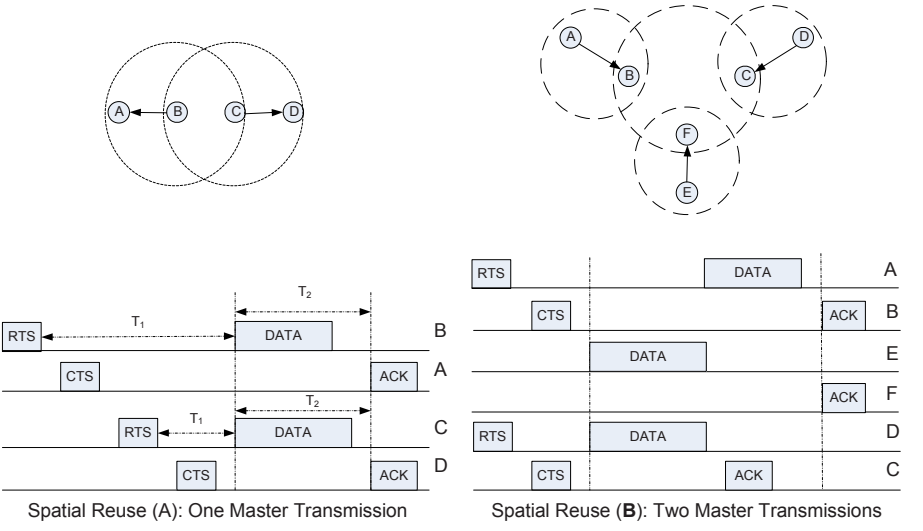


Fig. 2. Exploit Spatial Reuse

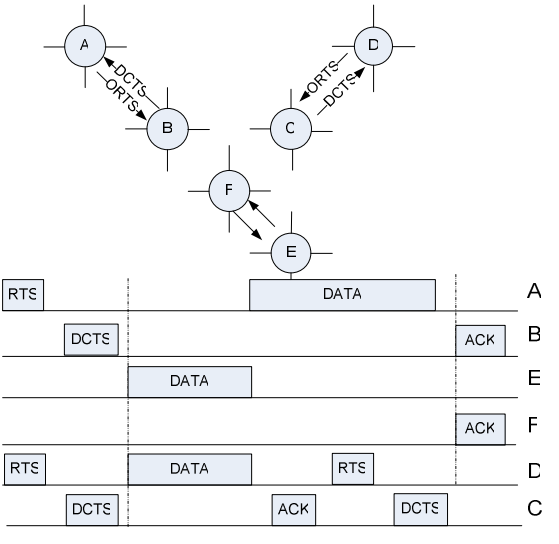


Fig. 3. Directional MAC Using Control window

node D initialize communication to node B and node C respectively. First, node A and node D send RTS packet omni-directionally. Node B and node C receive the RTS packet successfully, and determine the direction of node A and node D, then choose the beam to send the CTS packet directionally. After node A and node D receive the CTS packet, they can also determine the direction of node B and node C and choose the correct beam to send DATA. Within duration of control window, node E and node

F exchange the RTS/CTS packets. The data transmission from node E to node F align with data transmission from node D to node C, and ACK packet transmission from node F to node E align with ACK packet transmission from node B to node A. Before node F sends its ACK packet to node E, if node C has data send to node D, after node C receive RTS packet, it will send DCTS to node D. Thus node F will not receive DCTS from node D. In other word, this subsequent DCTS packet could not interfere with node F's reception of data. In a word, transmission equipped with directional antenna and control window, not only determine the position of nodes, but also improve the spatial reuse.

4 Performance Analysis

We evaluate the performance of our MAC protocol in compairison with IEEE 802.11 **[8]** using QualNet [5]. We design the simulation model with the grid topology. The 3 × N nodes are placed within a square area of length 1500 meters. We assume that the first row and the third row are the destination and the second row is source. There are N transmission column flows establish from source to destination and increase from left to right. The transmission range of the omni-directional antenna is 250m and that of the directional antenna is 500m. The data rate is 1Mbps. We assume node equipped with M=4 directional antennas because of grid topology.

We defined frame size is 2304 bytes because it is maximum MSDU (MAC Service Data Unit) on [8]. For the size of control window, as we know, if the length of control window is small, other pair of nodes cannot exchange their RTS/CTS completely, it results in poor spatial reuse. Whereas the length of control window is very big, after other pair of nodes exchanging their RTS/CTS completely, they are need to wait for a long time before exchanging DATA, so it results in poor utilization of the channel. So the size of the control window is made adaptive depending on the traffic in the network. Here, we define the size of the control window as a multiple of the time for a control packet exchange. A node that defines the control window choose a value for it as k × number of control packet exchange it heard in the previous window × time for a control packet exchange, where $1< k < 2$.

Fig. 4 shows the throughput of our MAC protocol which compares with IEEE 802.11 MAC protocol. For IEEE 802.11 MAC protocol, each node sends packets omni-directionally. In this situation, RTS/CTS reserve the use of channel and other neighbor nodes that overhear these packets must defer their transmission. Thus, N from 1 to 3, there is only one transmission can be allowed. Thus the performance of throughput is almost the same. When N is from 4 to 6, there are more ongoing transmissions at the same time. However, our MAC protocol which supports concurrent transmission get a higher performance of throughput than IEEE 802.11 MAC protocol, since equipped directional antenna and control window. With the increasing of number of N, there are more than two ongoing transmissions communicating with each other, so the performance of throughput grows. Therefore, our protocol realizes spatial reuse and prevents collision.

The performance of average end-to-end delay is evaluated in Fig. 5. The end-to-end delay is the time interval calculated from the instance a packet is handed down by the application layer at the source node, to the instance the packet is received by the

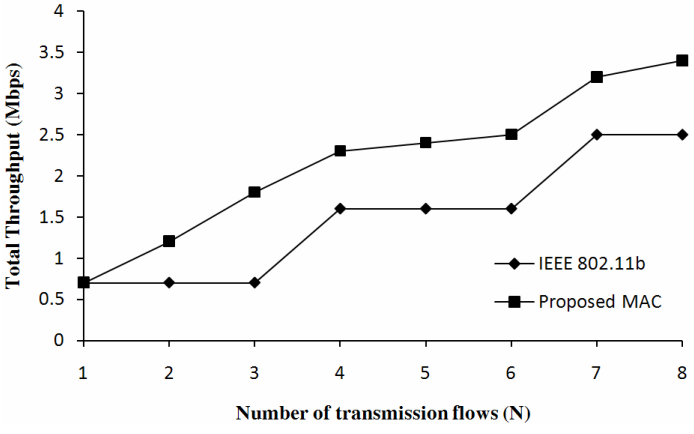


Fig. 4. Throughput in comparison with IEEE 802.11 MAC protocol

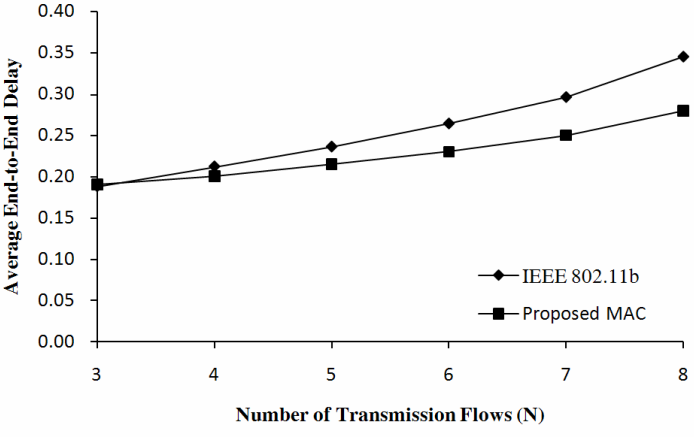


Fig. 5. Average End-to-End Delay

application layer at the destination node. We observe that the average of end-to end-delay of both of two protocols increase with increase in number of nodes. The delay of IEEE 802.11 MAC protocol increased quickly, while the increscent of our protocol is low.

5 Conclusions

In this paper we first discuss IEEE 802.11 MAC protocol which has prevented collision but limited to support for spatial reuse. Then we propose a new MAC protocol which address the position communication pairs by ORTS and DCTS and improve the performance of spatial reuse by a control window which set between the RTS/CTS and DATA/ACK packets. The results show the potential for significant

throughput improvement and average end-to-end delay decline. In subsequent work, performance studies on a larger and random network topology are needed to exhaustively show the performance features of our proposed protocol.

References

1. Chandra, A., Gummalla, V., Limb, J.: Wireless Medium Access Control Protocols. IEEE Communications, Surveys and tutorials 3 (2000)
2. Acharya, A., Misra, A., Bansal, S.: High-performance Architectures for IP-based multi-hop 802.11 networks. IEEE Wireless Communications 10(5), 22–28 (2003)
3. Choudhury, R., Yang, X., Ramanathan, R., Vaidya, N.: Using Directional Antennas for Medium Access Control in Ad Hoc Networks. In: International Conference on Mobile Computing and Networking, September, pp. 59–70 (2002)
4. Shankarkumar, K.Y., Vaidya, V.,, N.: Medium Access Control Protocols Using Directional Antennas in ad hoc Networks. In: Proceedings of IEEE Conference on Computer Communications, March, vol. 3, pp. 26–30 (2000)
5. QualNet user's Manual, `http://www.scalable-networks.com`
6. Takai, M., Martin, J., Ren, A., Bagrodia, R.: Directional Virtual Carrier Sensing for Directional Antennas in Mobile Ad Hoc Networks. In: Proc. ACM MobiHoc, June, pp. 183–193 (2002)
7. Nasipuri, A., Ye, S., You, J., Hiromoto, R.E.: A MAC Protocol for Mobile Ad-Hoc Networks Using Directional Antennas. In: Proceedings of IEEE Wireless Communication and Networking Conference, Chicago, IL, September, pp. 1214–1219 (2000)
8. IEEE Computer Society LAN MAN Standards Committee: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. In ANSI/IEEE Std. 802.11, 1999 Edition, the Institute of Electrical and Electronic Engineers, New York (1999)

Nonparametric Classification Based on Local Mean and Class Mean

Zeng Yong[1], Wang Bing[2], Zhao Liang[1], and Yu-Pu Yang[1]

[1] Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China
{zeng_yong,ypyang,zhaol}@sjtu.edu.cn
[2] School of Information and Electrical Engineering, Panzhihua University, Panzhihua 61700, China
wangb1009@163.com

Abstract. The *k*-nearest neighbor classification rule (*k*-NNR) is a very simple, yet powerful nonparametric classification method. As a variant of the *k*-NNR, a nonparametric classification method based on the local mean vector has achieved good classification performance. In this paper, a new variant of the *k*-NNR, a nonparametric classification method based on the local mean vector and the class mean vector has been proposed. Not only the information of the local mean of the *k* nearest neighbors of the unlabeled pattern in each individual class but also the knowledge of the ensemble mean of each individual class are taken into account in this new classification method. The proposed classification method is compared with the *k*-NNR, and the local mean-based nonparametric classification in terms of the classification error rate on the unknown patterns. Experimental results confirm the validity of this new classification approach.

Keywords: *k*-nearest neighbor classification rule (*k*-NNR), Local mean, Class mean, Cross-validation.

1 Introduction

As one of the most popular nonparametric techniques, the *k*-nearest neighbor classification rule (*k*-NNR) is widely used in pattern classification applications. According to this rule, an unclassified sample is assigned to the class represented by a majority of its *k* nearest neighbors in the training set. Cover and Hart have shown that, as the number *N* of samples and the number *k* of neighbors both tend to infinity in such a manner that $k/N \to 0$, the error rate of the *k*-NNR approaches the optimal Bayes error rate [1]. The *k*-NNR generally achieves good results when the available number of prototypes is large, relative to the intrinsic dimensionality of the data involved. However, in the finite sample case, the *k*-NNR is not guaranteed to be the optimal way of using the information contained in the neighborhood of unclassified pattern, which often leads to dramatic degradations of *k*-nearest neighbor (*k*-NN) classification accuracy. In addition to dramatic degradation of classification accuracy in small sample case, the classification performance of the *k*-NNR usually suffers from the existing outliers [3].That clearly explains the growing interest in finding variants of the *k*-NNR to improve the *k*-NN classification performance in small data set situations.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 593–600, 2008.
© Springer-Verlag Berlin Heidelberg 2008

As a variant of the *k*-NNR, a nonparametric classification method based on the local mean vector has been proposed by Mitani and Hamamoto [4]; this classification method not only overcomes the influence of the outliers but also can achieve good classification performance especially in small training sample size situations. That classification is equivalent to the nearest neighbor classification (1-NN) when the number of the neighbors in each individual class has r=1 and to the Euclidean distance (ED) classification when the number of the neighbors in each individual class equals the number of the samples in the corresponding class. Classifying with that classification method, it is easily noted that the number of the nearest neighbors used of the unlabeled pattern in each class is less than the number of the samples of the corresponding class in most situations. In other words, the classification method proposed by Mitani and Hamamoto utilizes the information of the local mean of the samples in each individual class more than the information of the corresponding class mean in most cases. It is natural to ask why not to design a classifier and such classifier utilizes the information of the local mean of the nearest neighbors of the test pattern in each individual class as well as the knowledge of the class mean of the corresponding class.

In pattern classification domain, the most important statistics-the sample mean and the sample covariance matrix constitute a compact description of the data in general. Matrix which is related to class separability, such as Bhattacharyya distance, scatter matrix, and Fisher's discriminant, is defined directly in terms of the mean and covariance matrix of the various classes. In the case that all the class mean vectors are not mutually same, to classify with the combination of information of the class mean vector and the local vector may be advisable. Motivated by the idea mentioned above, a nonparametric classification algorithm based on local mean and class mean has been proposed in this paper. This paper is organized as follows: the nonparametric classification algorithm based on local mean and class mean is presented in section 2. In section 3, the proposed classification algorithm is compared with the k-nearest neighbor classification rule (*k*-NNR), and the local mean-based nonparametric classification [4] in terms of the classification error rate on the unknown patterns. Section 4 concludes the paper and presents the future work.

2 The Proposed Nonparametric Classification Algorithm

2.1 The Proposed Nonparametric Classification Algorithm

For the N available labeled prototypes, let $N_1,...,N_M$ be the number of them belonging to the class $\omega_1,\omega_2,...,\omega_M$, respectively. Let $\boldsymbol{x}_j^{(1)},...,\boldsymbol{x}_j^{(r)}$ denote the r nearest neighbors in the j-th class prototypes of the unlabeled pattern $\boldsymbol{x}$ and $X_j=\{\boldsymbol{x}_j^i \mid i=1,...,N_j\}$ be a prototype set from class ω_j. Let $\boldsymbol{\mu}_j$ be the class mean vector for class ω_j:

$$\boldsymbol{\mu}_j=\frac{1}{N_j}\sum_{i=1}^{N_j}\boldsymbol{x}_j^i. \tag{1}$$

The proposed nonparametric classification is described as follows:

First, compute the local mean vector $\boldsymbol{y}_j$ using the r nearest neighbors in the j-th class prototypes of the unlabeled pattern $\boldsymbol{x}$:

$$\boldsymbol{y}_j = \frac{1}{r}\sum_{i=1}^{r} \boldsymbol{x}_j^{(i)}. \tag{2}$$

Calculate the distance d_j^l between the local mean vector $\boldsymbol{y}_j$ and the test pattern $\boldsymbol{x}$:

$$d_j^l = (\boldsymbol{x} - \boldsymbol{y}_j)^T (\boldsymbol{x} - \boldsymbol{y}_j). \tag{3}$$

Compute the distance d_j^c between the class mean vector $\boldsymbol{\mu}_j$ and the test pattern $\boldsymbol{x}$:

$$d_j^c = (\boldsymbol{x} - \boldsymbol{\mu}_j)^T (\boldsymbol{x} - \boldsymbol{\mu}_j). \tag{4}$$

Calculate the combining distance d_j :

$$d_j = d_j^l + w * d_j^c \quad 0 \le w \le 1. \tag{5}$$

Finally, classify the test pattern $\boldsymbol{x}$ into class ω_c if:

$$d_c = \min\{d_j\} \;\; j = 1, 2, ..., M \tag{6}$$

The parameter w is a reflection which the class mean vector exerts an influence on the classification effect, the greater the more obvious. It is obvious that the parameter w should be selected ranging from 0 to 1 because the local mean would exert a trivial influence on classification performance if the parameter w is greater than 1 and evidently the parameter w is not minus, and all experiments we have conducted have convinced it. The parameter w is selected by ranging from 1 to 0. In our experiments, a value selected from the following values is as the value of the parameter w:

$$w = 1.25^{-(i-1)} \;\; i = 1, ..., 41 \;\text{ or }\; w = 0. \tag{7}$$

The proposed nonparametric classification based on local mean and class mean is equivalent to the nonparametric classification based on local mean proposed by Mitani and Hamamoto [4] when $w = 0$. For the proposed nonparametric classification based on local mean and class mean, the key is how to get the suboptimal values about the parameter r, the number of nearest neighbors in each class, and the parameter w, the distance weight. Being similar to Mitani and Hamamoto [4], the optimization approach with the cross validation method is considered. The number of nearest neighbors in each class, the parameter r, and the distance weight, the parameter w can be set by cross-validation.

The suboptimal value of the parameter r is obtained by the *m-fold cross-validation* method [2] with the nonparametric classification based on local mean, and the *m-fold cross-validation* method used here is different from the cross validation method adopted by Mitani and Hamamoto [4]. The procedure for getting the suboptimal parameter r^* is as follows:

Step 1. Divide the *N* available total training samples into the training set and test set at random, each time with a different test set held out as a validation set.

Step 2. Classify with the nonparametric classification method based on local mean with the different parameter *r*.

Step 3. Get the error rate $e_i(r)$.

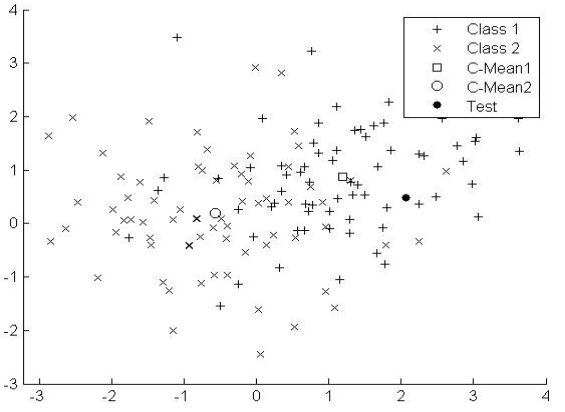


Fig. 1. The distribution of the samples and the indications of the corresponding class mean vectors and test pattern

Step 4. Repeat steps 1-3 *m* times to get the following:

$$E_{cv}(r)=\frac{1}{m}\sum_{i=1}^{m} e_i(r). \tag{8}$$

$E_{cv}(r)$ can be influenced by *m*. In the small training samples case, *m* equals the number of the total training samples, otherwise *m* is less than the number of the training samples. The suboptimal value of the parameter *r* is determined by

$$r^*=\arg\min\{E_{CV}(r)\}. \tag{9}$$

Finally, to get the suboptimal value of the parameter *w* by the *m-fold cross-validation* method [3] with the nonparametric classification based on local mean and class mean. This procedure is described as follows:

Step 1. Divide the *N* available total training samples into the training set and test set at random, each time with a different test set held out as a validation set.

Step 2. To classify by the nonparametric classification method based on local mean and class mean with the suboptimal parameter *r** and the different value of the parameter *w*.

Step 3. Get the error rate vector $\boldsymbol{e}_j(w,r^*)$.

Step 4. Repeat steps 1-3 *m* times to get the following:

$$\boldsymbol{E}_{cv}(w, r^*) = \frac{1}{m}\sum_{j=1}^{m} \boldsymbol{e}_j(w, r^*). \tag{10}$$

Being same to the procedure for getting the parameter r^*, $E_{cv}(w, r^*)$ can be influenced by m. In the small training samples case, m equals the number of the total training samples, otherwise m is less than the number of the training samples. In the procedure for getting the suboptimal parameter w^*, a value is selected for the parameter w in successive turns according to the equation (7). The suboptimal value of the parameter w is determined by

$$w^* = \arg\min\{\boldsymbol{E}_{cv}(w, r^*)\}. \tag{11}$$

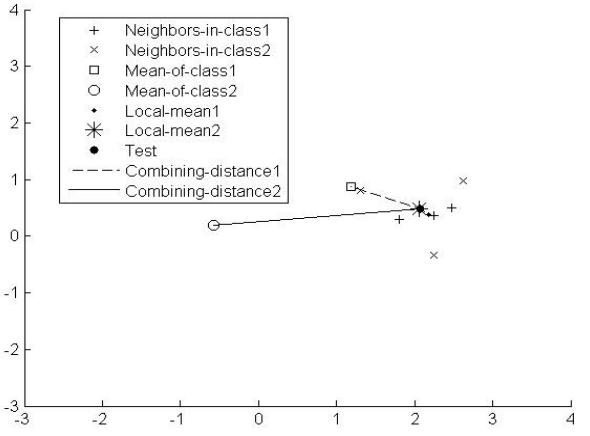


Fig. 2. Difference between the two nonparametric classification methods. For r = 3, the nonparametric classification method based on local mean assigns the test pattern into class 2 while the nonparametric classification method based on local mean and class mean assigns it, the class 1. In this special case, it is reasonable that the unlabeled pattern is labeled the class 1.

2.2 The Difference between the Two Nonparametric Classification Methods

Our proposed nonparametric classification based on local mean and class mean is very similar to the nonparametric classification based on local mean proposed by Mitani and Hamamoto [4] but different from it. The similarity between the two classification methods is using the information of the local mean of the nearest neighbors of the unlabeled pattern in each individual class to classify. The difference between the two classification methods is that the nonparametric classification based on local mean and class mean also uses the knowledge of the class mean of the training samples in each individual class to classify, and this difference can be illustrated through Fig. 1 and Fig. 2. Fig.1 gives the distribution of the samples and the indications of the corresponding class mean vectors and test pattern. Fig. 2 illustrates the difference between the two methods and the influence of the class mean vector exerting on the

classification performance. From Fig.1, we observe that the test pattern lies in the region which the samples from class ω_1 are dense, and it is rational to classify the test pattern into class ω_1. From Fig. 2 and in $r = 3$ case, the test point and the local mean point overlapped, and the test pattern is assigned into class ω_2 with the classification method proposed by Mitani and Hamamoto [4]. Just owing to influence of the class mean vector, the test pattern is classified into class ω_1 with the nonparametric classification method based on local mean and class mean, and it is obvious that this classification is reasonable.

3 Experimental Results

The classification performance of the proposed nonparametric classification method has been empirically assessed through the experiments on the several standard UCI Databases [5]. The local mean-based nonparametric classification method has been proposed by Mitani and Hamamoto, and they have compared their proposed method with other four nonparametric methods, 1-NN, *k*-NNR [1], Parzen [6], and ANN [7] in terms of the average error rate [4]. Their experimental results showed that their proposed classification method usually outperforms the other classification methods in terms of the error rate in most cases [4]. Our proposed classification method is an improvement of the classification method proposed by Mitani and Hamamoto, so it also possesses the classification attribute of the latter. Here, we only compared our proposed classification method with the classical *k*-NNR and the nonparametric classification method proposed by Mitani and Hamamoto [4]. For the sake of simplicity, we will refer to our proposed classification method as **LMCM**, refer to the *k*-nearest neighbor classification method as **KNN**, and refer to the nonparametric classification method proposed by Mitani and Hamamoto [4] as **LM** respectively later.

In our experiments, the distance metric used is the Euclidean distance. For the UCI standard data sets, only datasets with numeric features were selected. Some characteristics of the UCI standard data sets used are described in Table 1. For the data set **Letter** and **Thyroid**, the training set and test set are assigned in advance in earlier work on the same data set. For four other data sets, repeated experiments were done by randomly shuffling the whole dataset 100 times, and using the first 80% for the training and the rest for testing. We reported the average as result over these 100 repetitions and given the 95% confidence interval.

The experimental results on UCI standard data sets with three classification methods were shown in Table 2. Boldface is used to emphasize the best performing method(s) for each data set. For the data set **Letter** and **Thyroid**, the training set and test set are assigned in advance in earlier work on the same data set, so the corresponding optimal parameters of three classification methods conducting on the experiments can be gotten and were given in Table 2. From Table 2, it can be observed that the proposed nonparametric classification method based local mean and class mean (LMCM) performed well than two other classification methods except data set Iris. For the data set Iris, the classification method KNN can achieve a good performance than LM and LMCM.

Table 1. Some characteristics of the datasets used

Data	Features	Samples	Classes	Test set.
Letter	16	20,000	26	4,000 cases
Thyroid	21	7,200	3	3,428 cases
Pima	8	768	2	--
Wine	13	178	3	--
Iris	4	150	3	--
Glass	9	214	6	--

Table 2. Results on UCI standard data sets

Data		Error rate (100%)	95% Confidence interval (100%)	Optimal parameter
Glass	KNN	26.19	25.1~27.28	-
	LM	24.57	23.57~25.57	-
	LMCM	**23.45**	**22.46~24.44**	-
Letter	KNN	4.12	-	k=3
	LM	3.62	-	r=3
	LMCM	**3.6**	-	r=3 w=0.0024
Iris	KNN	**1.49**	**1.08~1.9**	-
	LM	1.76	1.32~2.2	-
	LMCM	1.76	1.32~2.2	-
Pima	KNN	22.57	22.04~23.1	-
	LM	22.86	22.33~23.39	-
	LMCM	**22.4**	**21.88~22.91**	-
Thyroid	KNN	6.33	-	k=5
	LM	5.92	-	r=5
	LMCM	**5.89**	-	r=5 w=0.0038
Wine	KNN	21.36	20.39~22.33	-
	LM	19.36	18.34~20.39	-
	LMCM	**18.56**	**17.51~19.6**	-

4 Conclusions

In this paper, a new nonparametric classification approach based on the local mean vector and the class mean vector is proposed. This new classification approach not only utilizes the information of the local mean of the nearest neighbors of the unclassified sample point but also utilizes the knowledge of class mean in each class. The experimental results show that this new classification nonparametric method usually performs better than the traditional k-nearest neighbor classification method, and is also superior to the local mean-based nonparametric classification method [4]. It should be pointed out that the classification performance is not improved in a particular situation that the class mean vectors are mutually equal. Two important statistics related to class discriminatory information, the class mean and the class covariance matrix, there is only the class mean being used to classify. In the case that each class mean is same, how to utilize the information of the class covariance matrix to classify should be further explored.

References

1. Cover, T.M., Hart, P.E.: Nearest Neighbor Pattern Classification. IEEE Transactions on Information Theory 13(1), 21–27 (1967)
2. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, 2nd edn. John Wiley & Sons, New York (2001)
3. Fukunaga, K.: Introduction to Statistical Pattern Recognition, 2nd edn. Academic Press, Boston (1990)
4. Mitani, Y., Hamamoto, Y.: A Local Mean-based Nonparametric Classifier. Pattern Recognition Letters 27(10), 1151–1159 (2006)
5. Asuncion, A., Newman, D.J.: UCI Machine Learning Repository. University of California, School of Information and Computer Science, Irvine (2007), http://www.ics.uci.edu/~mlearn/MLRepository.html
6. Jain, A.K., Ramaswami, M.D.: Classifier Design with Parzen Window. In: Gelsema, E.S., Kanal, L.N. (eds.) Pattern Recognition and Artificial Intelligence. Elsevier Science Publishers, North-Holland (1988)
7. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning internal representations by error propagation. In: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, ch. 8. MIT Press, Cambridge (1986)

Narrowband Jammer Excision in CDMA Using Particle Swarm Optimization

Imran Zaka[1], Habib ur Rehman[1], Muhammad Naeem[2], Syed Ismail Shah[3], and Jamil Ahmad[3]

[1] Center for Advanced Studies in Engineering (CASE), Islamabad, Pakistan
[2] Simon Fraser University, School of Engineering Science, Canada
[3] Iqra University Islamabad Campus, H-9, Islamabad, Pakistan
imran.zaka@gmail.com

Abstract. We formulate the narrowband jammer excision as an optimization problem and present a new algorithm that is based on Particle Swarm Optimization (PSO). In this paper particle positions represent the filter weights and particle velocity represents the updating increment of the weighting matrix. Simulation and numerical results show that PSO based algorithm is a feasible approach for jammer excision. It approaches the optimum performance in fewer iterations with lower complexity. The performance improvement is shown by plotting the Bit Error Rate (BER) as a function of Signal to Interference and Noise Ratio (SINR).

Keywords: Narrowband Interference, Jammer Excision, PSO, CDMA.

1 Introduction

In communication receivers, jammer excision techniques play a critical role in improving its performance. These techniques are efficient way to reduce the effect of jamming. These are employed at a pre-processing stage in the receiver. Noise in a communication channel can be intentional or un-intentional. Intentional noise is generally referred to as jamming and un-intentional as interference. Jamming is a procedure that attempts to block the reception of a desired signal by the intended receiver [1]. Generally it is a high power signal that occupies the same space, time slot or frequency spectrum as the desired signal, making reception by the intended receiver difficult or impossible. Jammer excision techniques are intended to counter this threat.

All jammer excision techniques try to achieve dimensional isolation of the jammer. These identify dimensionality of the jammer (space, time, frequency or combination of these) and then isolate it from that of the desired signal. These techniques result in some degradation of the desired signal for example suppressing jammer frequencies will also lose desired signal that occupies the same frequency band.

These techniques often need to be time varying because of the dynamic or changing nature of the jamming signal and the channel. The design of an optimum excision filter requires a priori information about the statistics of the data to be processed.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 601–609, 2008.
© Springer-Verlag Berlin Heidelberg 2008

There exist many diverse techniques for jammer excision, which includes notched filters [2], direct jammer synthesis [3], amplitude domain processing and adaptive antennas. These techniques are classified broadly as adaptive notch filtering, decision feedback, adaptive analog to digital conversion and non-linear techniques [4].

The methods for suppressing interference/jamming of various types in Direct Sequence (DS) and Frequency Hopped (FH) spread spectrum systems is discussed in [5] and it has been concluded that linear and nonlinear filtering methods are particularly effective in suppressing continuous-time narrowband interference. In [6-7] performance improvement has been shown by the use of transversal filters in DS spread spectrum system in the presence of Narrowband Interference (NBI). Various methods have been proposed for NBI suppression in [8-11]. Adaptive algorithms are proposed for estimation and suppression of NBI in DS spread spectrum system in [12]. Bijjani and Das have applied linear and non linear neural network filters for suppression of NBI in DS spread spectrum system [13]. Higher order statistics and Genetic Algorithm (GA) has been used to reject NBI and has faster convergence than Least Mean Squares (LMS) algorithm [14]. Recently a new evolutionary computation technique called Particle Swarm Optimization (PSO) has been proposed and introduced [15-16]. PSO has been motivated by the behavior of organisms such as fish schools and bird flocks. PSO is characterized as a simple in concept, easy to implement and computationally efficient. Unlike other heuristic techniques PSO has a flexible and well-balanced mechanism to enhance and adapt to global and local explorations abilities. There are wide varieties of problems that have been solved using PSO [17-19]. PSO has been applied for achieving global optimization in non-linear and recursive adaptive filter structures [20-21].

In this paper, the problem of jammer excision in Direct Sequence-Code Division Multiple Access (DS-CDMA) system is considered. The jammer excision is formulated as an optimization problem and then solved by PSO based approach. Rest of the paper is organized as follows. Next section describes the model for the DS-CDMA system with narrow band jammer. Section 3 presents Wiener filter for excision in CDMA. Section 4 provides an overview of PSO. Section 5 describes the algorithm of PSO for jammer excision. Computational complexity and simulation results are considered in section 6 and 7 respectively and section 8 concludes the paper.

2 System Model

We consider a synchronous CDMA system. CDMA system with jammer excision is shown in figure 1 for a single user y. Consider the k-th user transmitting Binary Phase Shift Keying (BPSK) symbols. The users are separated by Pseudo Noise (PN) spreading sequences of length *L*. The m-th symbol of k-th user is spread over L chips using a unit energy spreading sequence $c_k = \{c_k(1), c_k(2), c_k(3), ..., c_k(L)\}$, where $c_k(l) \in \{\pm 1/\sqrt{L}\}$, $l = 1, 2, ..., L$. The complex equivalent low pass transmitted signal for k-th user can be written as:

$$s_k(t) = \sum_{i=-\infty}^{\infty} \sum_{l=0}^{L-1} a_k(i)\left(c_k(l)\, p(t - lT_c - iT_s\right) \tag{1}$$

where $a_k(i)$ and $c_k(l)$ are the i-th information bit and the l-th chip of the spreading code of k-th user. *L* is the length of spreading code and is called the processing gain, $L = T_s / T_c$, T_s =1/(symbol rate) is the symbol duration and T_c, the chip duration. In (2), *p*(*t*) is a pulse satisfying the following relations:

$$p(t) = \begin{cases} 1, & 0 \le t \le T_c \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

The received signal in the presence of a jamming signal and additive white Gaussian noise (AWGN) can be written as:

$$r(t) = \sum_{k=1}^{K} \sum_{i=1}^{\infty} \sum_{l=0}^{L-1} a_k(i)\, c_k(l)\, p(t - kT_c - iT_s) + J(t) + n(t) \tag{3}$$

where K is the total number of users. *n(t)* is the AWGN and *J(t)* is the jamming signal. J(t) can be written as sum of sinusoids:

$$J(t) = \sum_{m=1}^{M} A_m \sin(2\pi f_m t + \varphi_m) \tag{4}$$

Where *Am*, *fm* and *φm* are amplitude, frequency and phase of mth sinusoid.

The received signal r(t) is then sampled to get r(n) and convolved with a discrete excision filter. The output of the filter with N filter coefficients/weights w is denoted by y(n) and is given by:

$$y(n) = \sum_{j=1}^{N} r(n-j) w(j) \tag{5}$$

The received signal is then passed through a bank of K correlators or matched filter prior to decision stage. This involves multiplication with users spreading code and averaging it over symbol duration T_s, output of the k^{th} correlator at receiver is given by

$$u_k(n) = \frac{1}{L} \sum_{i=1}^{T_s} \sum_{l=0}^{L-1} y(i)\, c_k(l) \tag{6}$$

The bit is detected by a conventional single user detector by just determining the sign of the correlator output. Let the detected bit be $\hat{a}_k$ using a matched filter the detection is made as

$$\hat{a}_k(n) = sign(\text{Real}(u_k(n))) \tag{7}$$

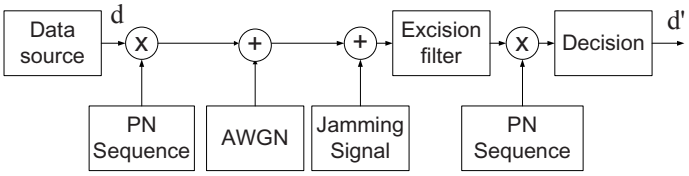


Fig. 1. Block diagram of a CDMA System with an Excision Filter

3 Wiener Filtering for Jammer Excision in CDMA

Wiener filter reduces the effect of jamming by comparison with the desired noiseless signal. Pilot sequence is transmitted for designing of filter and is known at receiver. Performance criterion of Wiener filter is minimum mean square error. The Mean Square Error (MSE) is defined as:

$$E\left[e^2(n)\right] = \frac{1}{N}\left[\sum_{k=1}^{N}\left(d_{pilot}(n-k) - y_k(n)\right)^2\right] \tag{8}$$

The Wiener filter is designed to achieve an output close to the desired signal d_{pilot} by finding the optimum filter coefficients that minimize the MSE between the pilot data and filtered signal, which can be stated as:

$$\mathrm{w}_{opt} = \arg\min\ E\left\{e^2[n]\right\} \tag{9}$$

The Wiener filter coefficients wopt are given by:

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{P}\,. \tag{10}$$

Where R is the autocorrelation matrix of the received pilot signal and P, the cross correlation matrix between the received signal and the pilot signal. Jammer free signal is achieved using this wopt in equation (5). The output of the filter is further processed for detection.

The Wiener filter solution is based on the assumptions that the signal and the noise are stationary linear stochastic processes with known autocorrelation and cross correlation. In practice the exact statistics (i.e. R and P) are not known, needed to compute the optimal Weiner filter hence degrading the performance. Larger size of R and P is required for more accurate estimates of correlation values resulting in large and better wopt. The large sizes of R, P and wopt are too expensive in many applications (e.g. real time communication). Efficient methods are required for calculation of matrix inverse (R^{-1}).

We propose in this paper jamming excision based on particle swarm optimization that does not require the known statistics of the signal and the noise. PSO based method also alleviate cumbersome calculations for finding inverse of matrix.

4 Overview of Particle Swarm Optimization

Particle swarm optimization is a robust optimization technique developed by Eberhart and Kennedy in 1995 [15-16] based on the movement and intelligence of swarms. It applies the concept of social interaction to problem solving using a number of agents (particles) that constitute a swarm moving around in the search space looking for the best solution. Each particle is treated as a point in an N-dimensional space which adjusts its "flying" according to its own flying experience as well as the flying experience of other particles keeping track of its coordinates in the solution space which are associated with the best solution (fitness) that has achieved so far by that particle. This value is called personal best or *pbest*. Another best value that is tracked by the PSO is the best value obtained so far by any particle in the neighborhood of that particle. This value is called *gbest*. The basic concept of PSO lies in accelerating each particle towards its *pbest* and the *gbest* locations, with a random weighted acceleration in each iteration. After finding the two best values, the particle updates its velocity and position equations.

$$\mathbf{v}_i^{k+1} = \boldsymbol{\varphi}\, \mathbf{v}_i^k + \boldsymbol{\alpha}_1 \boldsymbol{\gamma}_{1i} \left(\mathbf{p}_i - \mathbf{x}_i^k \right) + \boldsymbol{\alpha}_2 \boldsymbol{\gamma}_{2i} \left(\mathbf{g} - \mathbf{x}_i^k \right) \tag{11}$$

$$\mathrm{x}_i^{k+1} = \mathrm{x}_i^k + \mathrm{v}_i^{k+1} \tag{12}$$

Where i is particle index, k is the iteration index, v is the velocity of the particle, x is the position of the particle, p is the pbest found by the particle, g is the gbest and γ1i, γ2i, are random numbers in the interval [0,1]. Two independent random numbers are used to stochastically vary the relative pull of gbest and pbest. In (11) α1,and α2 are the acceleration constants used to scale the contribution of cognitive and social elements and it can also be termed as learning factors and φ is the inertia function. The parameter φ is very important in determining the type of trajectory the particle travels. A large inertia weight facilitates the global exploration while with a smaller one the particle is more intended to do local exploration.

A proper choice of inertia weight provides the balance between the global and local exploration ability of the swarm. Experimental results suggest that it is better to initially set the inertia to be a large value and then gradually decrease its value to obtain the refined solution [22]. The velocity is applied for a given time-step, and the particle moves to the next position.

5 PSO for Jammer Excision in CDMA

In this paper, PSO is applied for jammer excision of CDMA signal to minimize the Mean Square Error (MSE) or cost function (8) between the pilot data and filtered signal. Particle position **w** represents the detector weights and particle velocity represents the updating increment Δ**w** in the weight matrix i.e

$$\Delta \mathbf{w}_i^{k+1} = \boldsymbol{\varphi}\, \Delta \mathbf{w}_i^k + \boldsymbol{\alpha}_1 \boldsymbol{\gamma}_{1i} \left(\mathbf{p}_i - \mathbf{w}_i^k \right) + \boldsymbol{\alpha}_2 \boldsymbol{\gamma}_{2i} \left(g - \mathbf{w}_i^k \right) \tag{13}$$

$$\mathbf{w}_i^{k+1} = \mathbf{w}_i^k + \Delta\mathbf{w}_i^{k+1} \tag{14}$$

First the initial population is generated with random w and Δw of dimensions of dimensions S×N, where S represents the swarm size and N is the filter length. The current searching point of each agent is set to pbest. The best evaluated value of pbest is set to gbest and agent number with best value is stored. Each particle evaluates the cost function (8). If this value is less than the current pbest of the agent, the pbest is replaced by the current value. If the best value of pbest is less than the current gbest, the gbest is replaced by the best value and the agent number with its value is stored. The current searching point of each agent is changed using (13) and (14). This process will continue until the termination criterion is finally met. PSO algorithm for jammer excision in CDMA is described in the following para.

PSO Algorithm

Initialize particles with initial weight vector $\mathbf{W}$ and increment vector $\Delta\mathbf{W}$ with dimensions $S\times N$. Let $\mathbf{w}_s$ represent the s^{th} row of $\mathbf{W}$

```
for   i =1: iterations
      for   s =1:S
            evaluate fitness function (8) for s^th row of W
            if (fitness(w_s) < fitness(w_pbest^(s)))
                  w_pbest^(s) = w_s
            end if
      end
            if(fitness_{1≤s≤S}(w_pbest^(s)) < fitness(w_gbest))
            w_gbest = w_{arg min_{1≤s≤S}(fitness(w_pbest^(s)))}
            end if

      update W and ΔW equations (13) and (14)
end
      w_gbest is the solution
```

6 Computational Complexity and Implementation Issues

The computation of Wiener filter coefficients, involve calculation of inverse of the correlation matrix. The computational complexity of inverting an n×n matrix by Gaussian Elimination leads to $o(n^3)$ [23-24]. Although Gaussian Elimination method is not optimal and there exists a method (Strassen's method) that requires only $o(n^{\log_2(7)}) = o(n^{2.807})$ operations for a general matrix. But the programming of Strassen's algorithm is so awkward, and often Gaussian Elimination is still the preferred method. The computational complexity of PSO based algorithm is derived from (11) and (12). It can be easily shown that complexity of PSO based algorithm is o(n) and we can say that PSO based algorithm is less complex than the Wiener filter.

One of the key advantages of PSO is the ease of implementation due to simplicity. Unlike computationally expensive matrix inversions, PSO merely consists of two straightforward equations (11) and (12) for weight updates and simple decision loop to update the pbest and gbest.

7 Simulation Results

In this section, we present the simulation results to show the performance. In these simulations, number of users is 16, processing gain is 64, filter length is 32, swarm size is 32, and number of iterations for PSO algorithm is 75. Fig. 2a shows the average BER performance of the system with varying bit energy and keeping the interference power constant. PSO based excision filter achieves the performance as that of a Wiener filter (optimal) with less computational complexity and ease of implementation. Performance is also shown for no excision filter and without jammer (NBI) cases in order to provide reference. Jammer excision filters provide mitigation against the jammer as evident when their performance is compared with the case of without mitigation. Increasing the bit energy has almost no effect on performance in the without mitigation case.

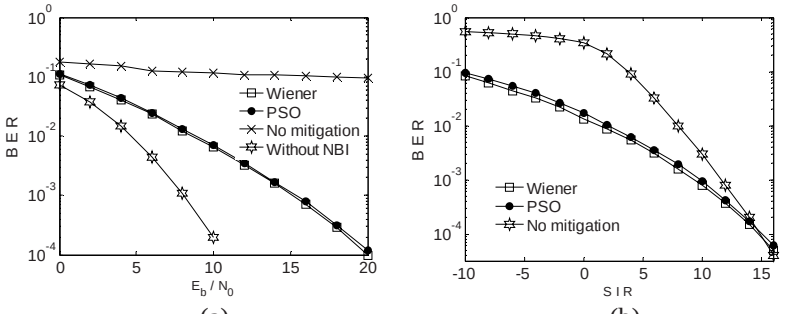


Fig. 2. BER performance. (a) Interference power being kept constant and varying bit energy. (b) Noise power being kept constant and varying interference power.

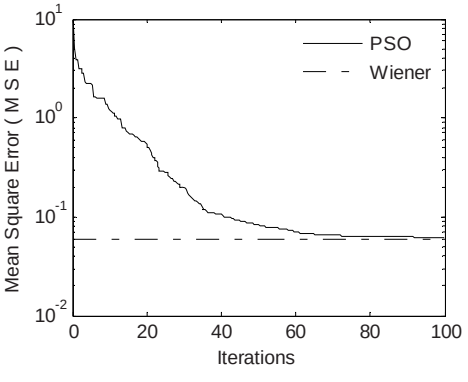


Fig. 3. Convergence of fitness or objective function for PSO

Fig.2b shows the BER performance of the system with constant noise power and varying interference power. It can be seen that PSO achieves the optimal Wiener filter performance. Mitigation against jamming is evident when performance of excision

filter is compared with no mitigation scenario. Excision filters provide more than 12 dBs of gain when their BER performance is observed at very low Signal to Interference Ratio (SIR) values (i.e. in the presence of high power jammer), which is usually the case in a jamming scenario. Some performance degradation is observed only at relatively high SIR values as excision of a low power jammer also removes part of the signal of interest. In such a scenario a SIR threshold is decided beyond which excision filter is turned off and received signal is directly used for detection. In this case SIR of 15dB is the threshold point. The convergence rate of objective function (8) with the number of generations or iterations is shown in fig. 3. These results have been obtained by averaging it over many iterations. It can be seen that PSO converges after about 75 iterations.

8 Conclusions

A PSO based jammer excision algorithm is proposed. Unlike conventional PSO algorithms, the proposed PSO based jammer excision algorithm uses filter weights as particle positions and particle velocity as updating increment/decrement of the weighting matrix. The weighting matrix can be found in finite iterations. The results are compared with the optimum Wiener filter. Wiener filter involves the calculation of inverse of a large matrix whose size depends upon the size of weighting matrix. Consequently complexity increases exponentially with the increase in size of weighting matrix. PSO based algorithm reduces the complexity and drastically simplifies implementation while giving the same optimal BER performance.

References

1. Papandreou-Suppapola, A.: Applications in Time-Frequency Signal Processing. CRC press, Tempe, Arizona (2003)
2. Amin, M., Wang, C., Lindsey, A.: Optimum Interference Mitigation in Spread Spectrum Communications using Open Loop Adaptive Filters. IEEE Transactions on Signal Processing 47(7), 1966–1976 (1999)
3. Lach, S., Amin, M.G., Lindsey, A.: Broadband Non-stationary Interference Excision in Spread Spectrum Communication System using Time Frequency Synthesis Techniques. IEEE journal on selected Areas in communications (1999)
4. Laster, J.D., Reed, J.H.: Interference Rejection in Digital wireless Communications. IEEE Signal Processing Magazine 14(3), 37–62 (1997)
5. Proakis, J.G.: Interference suppression in spread spectrum systems. In: Proceedings of 4th IEEE International Symposium on Spread Spectrum Techniques and Applications, vol. 1, pp. 259–266 (1996)
6. Li, L.M., Milstein, L.B.: Rejection of narrowband interference in PN spread spectrum systems using transversal filters. IEEE Trans. Communications COM-30, 925–928 (1982)
7. Theodoridis, S., Kalouptsidis, N., Proakis, J., Koyas, G.: Interference Rejection in PN Spread-Spectrum Systems with LS Linear Phase FIR Filters. IEEE Transactions on Communications 37(9) (1989)
8. Medley, M.J., Saulnier, G.J., Das, P.K.: Narrow-Band Interference Excision in Spread Spectrum Systems Using Lapped Transforms. IEEE Trans. Communications 45(11) (1997)

9. Shin, D.C., Nikias, C.L.: Adaptive Interference Canceller for Narrowband and Wideband Interference Using Higher Order Statistics. IEEE Trans. Signal Processing 42(10), 2715–2728 (1994)
10. Rusch, L.A., Poor, H.V.: Narrowband Interference Suppression in CDMA Spread Spectrum Communications. IEEE Trans. Commun. 42(234) (1994)
11. Duklc, H.L., DobrosavlJevic, Z.S., StoJanovlc, Z.D., Stojanovlc, I.S. (eds.): Rejection Of Narrowband Interference in DS Spread Spectrum Systems Using Two-Stage Decision Feedback Filters. IEEE, Los Alamitos (1994)
12. Ketchum, J.W., Proakis, J.G.: Adaptive Algorithms for: Estimating and Suppressing Narrow-Band Interference in PN Spread-Spectrum Systems. IEEE Transactions on Communications [legacy, pre - 1988] 30(5), part 2, 913–924 (1982)
13. Bijjani, R., Das, P.K.: Rejection Of Narrowband Interference In PN Spread-Spectrum Systems Using Neural Networks. In: Global Telecommunications Conference, and Exhibition. Communications: Connecting the Future, GLOBECOM 1990, vol. 2, pp. 1037–1041. IEEE, Los Alamitos (1990)
14. Taijie, L., Guangrui, H., Qing, G.: Interference Rejection in Direct-Sequence Spread Spectrum Communication Systems Based on Higher-Order Statistics and Genetic Algorithm. In: Proceedings of 5th International Conference on Signal Processing, vol. 3, pp. 1782–1785 (2000)
15. Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. In: Proceedings of International Conference on Neural Networks, Perth Australia, pp. 1942–1948. IEEE Service Center, Piscataway (1995)
16. Kennedy, J., Eberhart, R., Shi, Y.: Swarm Intelligence. Morgan Kaufmann Academic Press, San Francisco (2001)
17. Kennedy, J.: The Particle Swarm: Social Adaptation of Knowledge. In: Proceedings of the IEEE International Conference on Evolutionary Computation, pp. 303–308 (1997)
18. Kennedy, J., Eberhart, R.C.: A binary version of particle swarm optimization algorithm. In: Proceedings of Systems Man and Cybernetics (SMC 1997), pp. 4104–4105 (1997)
19. Yoshida, H., Fukuyama, Y., Kwata, K., Takayama, S., Nakanishi, Y.: A particle swarm optimization for reactive power and voltage control considering voltage security assessment. IEEE Transaction on power systems 15 (2001)
20. Krusienski, D.J., Jenkins, W.K.: Adaptive Filtering via Particle Swarm Optimization. In: Proc. of the 37th Asilomar Conf. on Signals, Systems, and Computers (2003)
21. Krusienski, D.J., Jenkins, W.K.: Particle Swarm Optimization for Adaptive IIR Filter Structures. In: Proc. of the 2004 Congress on Evolutionary Computation (2004)
22. Shi, Y., Eberhart, R.: Empirical Study of Particle Swarm Optimization. In: Proceedings of the Congress on Evolutionary Computation, pp. 1945–1950 (1999)
23. Higham, N.J.: Accuracy and stability of numerical algorithms. SIAM, USA (1996)
24. Householder, A.S.: The theory of matrices in numerical analysis. Dover Publications, USA (1974)

DHT-Based Mobile Service Discovery Protocol for Mobile Ad Hoc Networks*

Eunyoung Kang, Moon Jeong Kim, Eunju Lee, and Ungmo Kim

School of Computer Engineering, Sungkyunkwan University,
440-776, Suwon, Korea
{eykang,tops,eunju,umkim}@ece.skku.ac.kr

Abstract. Service discovery to search for a service available in a mobile ad-hoc network is an important issue. Although mobile computing technologies grow ever more powerful and accessible, MANET, consisting of mobile devices without any fixed infrastructure, has such features as high mobility and resource constraints. In this paper, we propose an effective service discovery protocol which is based on the concept of peer-to-peer caching of service information and DHT-based forwarding of service requests to solve these problems. Neighboring information, such as services and power is shared not only through logical neighbors but also through physically nearby neighboring nodes. Our protocol is that physical hop counts and the number of messages exchanged have been significantly reduced, since it does not require a central lookup server and does not rely on multicasting and flooding. The results of the simulation show that our proposed scheme works very well on dynamic MANET.

Keywords: Service discovery, Peer-to-Peer System, MANET, DHT, Structured.

1 Introduction

A mobile ad-hoc network (MANET) autonomously composed of mobile nodes independent of the existing wired networks or base stations has recently attracted substantial interest from industrial or research groups. Because it lacks infrastructure support, each node acts as a router, forwarding data packets for other nodes [1]. In accordance with these trends, such mobile devices as PDAs, handheld devices and notebook computers have rapidly evolved. Due to the development of those mobile devices increasing user demand, file sharing or service discovery is emerging as an important issue.

With regard to file sharing or service discovery, there have been lots of researches on structured/unstructured wired network P2P systems. Examples of representative unstructured P2P systems include Gnutella [2] and KaZaA [3]. Because Gnutella uses

* This work was supported in part by the MKE(Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the IITA(Institute of Information Technology Advancement, IITA-2008-C1090-0801-0028) and by Foundation of ubiquitous computing and networking project (UCN) Project, the Ministry of Knowledge Economy(MKE) 21st Century Frontier R&D Program in Korea and a result of subproject UCN 08B3-B1-10M.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 610–619, 2008.
© Springer-Verlag Berlin Heidelberg 2008

centralized directory services, it can easily find the location of each directory. But it causes a central directory server to produce a number of query messages, incurring bottlenecks. On the other hand, such structured P2P systems as Chord [4], CAN [5], and Pastry [6] are based on distributed hash tables (DHTs). A DHT-based overlay network makes out a table for searching by using a file and a key value obtained as a result of applying the file to its hash function. It's effective for peers involving themselves in a network to distribute and save the whole file information for searching because such searching is based on the key value corresponding to the file.

Notwithstanding advantages as demonstrated above, cases of applying those technologies to a mobile ad-hoc network are very few. It is difficult to apply P2P applications based on a fixed network environment to ad-hoc networks different from those fixed wired networks because their nodes are free to move. Although mobile computing technologies get more and more powerful and accessible than ever, mobile devices has a lower level of processing capacity and uses batteries with limits in their power. They consume lots of power when they exchange messages with one another. In this sense, it is needed to reduce costs for P2P applications with mobile devices in a wireless network. An ad-hoc network would enable reducing the consumption of energies, which is followed by the transmission of messages, and saving wireless bandwidth needed for such transmission, by lessening the number of query messages among P2P applications in the network.

To solve these problems, this study proposes an improved service discovery protocol which is based on the concept of peer-to-peer caching of service information and DHT-based forwarding of service requests. First, it listen service information from neighbor node and store service information in own local service directory for the purpose of caching service information. Second, it utilizes DHT to efficiently cache service information in MANET. The use of DHT not only reduces the message overhead, but also provides a useful guide which tells nodes where to cache their service information. We can reduce the number of message hops by maintaining additional information in a physical neighbor and information learned from nodes 2~3 hops away, and discovering possible shortcuts.

The remainder of this paper is organized as follows: in Section 2, the existing various approaches for service discovery are covered; Section 3 gives an explanation to the proposed service discovery architecture; an analysis of the architecture and its validity are exploited in Section 4; Finally, Section 5 concludes this paper, proposing the future study.

2 Existing Approaches for Service Discovery

Service discovery protocols have been designed for static networks including the Internet. Such well-known service discovery mechanisms as Service Discovery Protocol (SLP) [8], UDDI [9], and Jini [10] adopt an approach supporting central directory, while KaZaA, UPnP [11], SSDP, and JXTA are mechanisms supporting flooding. However, for mobile nodes to use services on a mobile ad-hoc network without a fixed infra architecture or central server, they need a service discovery mechanism of distributed architecture provided through collaborations among each

mobile node. Flooding approach creating lots of query messages produces a great deal of traffic when the size of a network grows larger, incurring lots of costs.

Service discovery protocols on a multi-hop ad-hoc network include Salutation [13], Konak [14], and Allia [15]. Salutation is a cooperation architecture developed by Salutation Consortium to enable communications among various development platforms of electric home appliances and devices and to provide connectivity and mobility in a distributed environment. Konak is an approach in which mobile nodes consist of P2P models in a multi-hop ad-hoc environment each of which can host its own local services, deliver them through a multi-http server and send queries to a network to identify possible services provided by another node. Allia is an agent-based service discovery approach in a mobile ad-hoc network environment, serving as a service discovery framework for mobile ad-hoc networks based on P2P cashing and policy. Each mobile node has its own agent platform.

Recently, there have been lots of studies using DHT-based overlay to share resources in an ad-hoc network [7]. In reference literature using similar synergies with an ad-hoc network and a P2P network spontaneously systematized and distributed, architectures of combining P2P and mobile applications was proposed. Pointing out logical vs. physical routing, they consider “combination” an issue sensitive to them. Of references, [16] and [17] show that costs of maintaining a strict overlay structure act as a material obstacle due to its big dynamicness and resource restrictiveness. In [16], you can see that conditions form overlay structure has been aggravated. On-demand structure and routing are proposed in [17]. Cell Hash Routing (CHR) [18] is characterized ad-hoc DHT. CHR uses location information by building DHT of clusters in stead of systematizing individual nodes in overlay. Routing among clusters is carried out in accordance with position-based routing of GPSR routing algorithm. Main restrictions to this approach are that nodes are not made addressable. Instead, they are made addressable only through clusters.

3 Proposed Service Discovery Architecture

3.1 Overview

Fig. 1 shows components included in a single mobile device and the proposed service discovery architecture. The service discovery architecture is composed of 3 layers: application layer, service management layer and network communication layer.

Application layer provides users with applications such as audio-video player, event alarm, network storage, etc. Service management layer provides services related to discovery. To register information about services provided from neighboring nodes on a network, cache and cache manager are used. Service provider stores its own services at LSD (Local Service Directory), periodically advertising them to its neighboring nodes. All devices on networks store information listened in their local caches for a given period and advertisement messages listened from a service provider. When a user wants to find out a service, service discoverers begin with searching LSD, and NSC (Neighbor Service Cache) in their local caches. If a requested service has not been found out in local cache, a service request message is sent to forwarding manager to start service discovery.

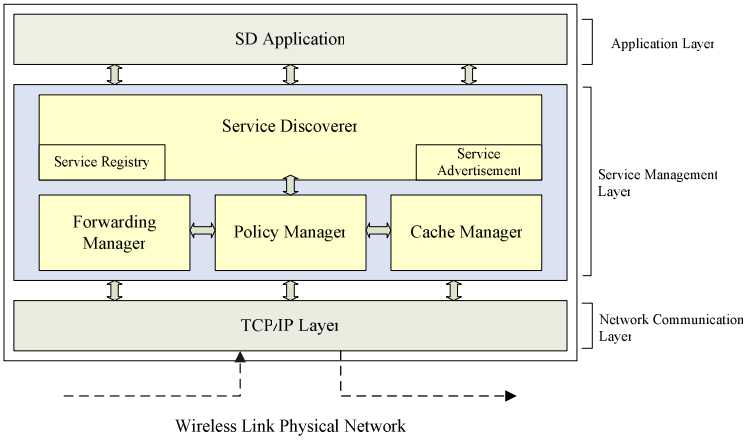


Fig. 1. Service Discovery Architecture

Forwarding manager plays a role in carrying service advertisement messages and request messages. Service discoverer picks out the shortest shortcut of neighboring nodes, logical nodes and any other nodes on paths passed by the by using service discovery protocol proposed in this paper to carry a request message. Policy manager plays a role in controlling nodes through their current policies which describe service advertisement preference or replacement strategy or refresh rate, TTL(time-to-live).

4 Proposed Service Discovery Protocol

4.1 Cooperative Neighbor Caching of Service Information

Each node on a network saves and manages information listened to at their local cache for a specified period. We call this cache "Neighbor Service Cache (NSC)," which contains such basic information, as about node connectivity degree, power information, IP address, service description, and timestamp. Each node as a service provider acts as not only a server but also a client when it requests for a needed service. We use AODV well-known as a routing protocol in a mobile ad-hoc network. Hello messages of AODV make a periodical one-hop broadcast for neighboring nodes. To identify information about the neighboring nodes of each node, we piggyback onto hello messages information, as about its id, service name, degree and power. Therefore, we do not use separate messages to obtain information about neighboring nodes. We produce a NSC to find out the existence of and information about neighboring nodes, based on hello messages obtained from them. This information is saved in NSC. Nodes newly joining in a network inform their own existence by making broadcast to their neighboring nodes via hello messages. If there is not any hello message got from those neighboring nodes for a specified period, they are beyond their transmission range or have left the network. We can reduce the number of message hops used because it is possible to maintain visibility of nodes 2~3 hops

away and additional information from physical neighbors and find out a possible shortcut.

4.2 DHT Utilizing Topology Information

The use of DHT not only further reduces the message overhead, but also provides a definite guide which tells nodes where to search available services. Each service description (SD) makes key using hashing and each node stores and maintains a (Key, SD) pairs on overlay routing table. Service advertisement (SA) is performed by the insertion of (Key, SD) to the node that takes charge of such area. Service discovery is performed by the lookup process on the overlay. When a node receives a service request message, overlay routing table is referred to and the request message travels along a path built from the routing table. The logical overlay routing does not consider the physical topology at all, so unnecessarily longer paths often introduce the inefficiency and waste of limited resource of MANET devices. But we do not use only long-range logical neighbors incurring a lot of costs. Instead, we use information of physically closer neighboring nodes to quickly reach a destination on an overlay network so that service request can return to its requester without reaching service provider through information cached in intermediate nodes.

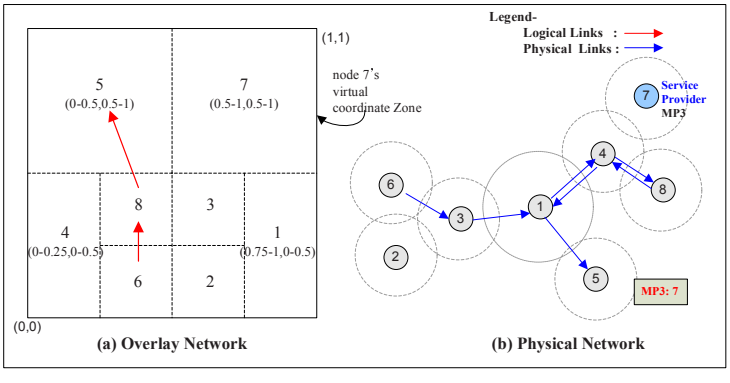


Fig. 2. Overlay Network & Physical Network

Fig. 2 shows a DHT-based overlay network and physical network. Let us suppose that node 7 is a service provider providing "abc.mp3" services, and that node 6 is a service requester requesting for those services. Figure 2 (a) represents overlay network mechanism for CAN using DHT. Figure 2 (b) represents a real physical network. If node 7, as a service provider, gets a value (0.4, 0.8) which is derived through hashing a service name called "abc.mp3," available services are also registered with DHT table of node 5 in charge of the same value in DHT. Node 6 knows through the value of key (0.4, 0.8) derived out by hashing "abc.mp3" to search available services that there is their registration information in node 5. Node 6 sends service request message with a (Key, SD) pairs. Node 6 searches for a path leading to node 5

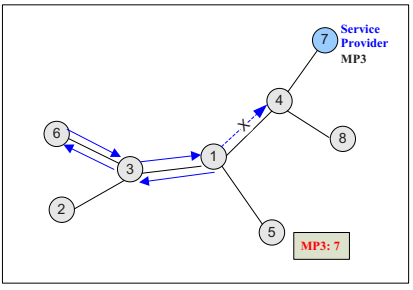


Fig. 3. Example using of physical neighbor's information

using DHT routing table to go through a routing path of n6-n3-n1-n4-n8-n4-n1-n5. However, DHT-based overlay routing is transmitted to logical neighbors, and independent of physical topology; that is, a physical network is not considered.

Fig.3 shows proposed example. We use information of neighboring nodes collected through a physical network to quickly reach a destination on an overlay network. Since service request message knows information about node 5 by using information cached at neighbor node 1, it is possible to send response messages to their service requester without reaching the destination. Therefore, the routing path is as follows: n6-n3-n1. It meant that, since the number of hops ranging from service request node to service response node is so shortened as to find out the shortest shortcut, a consumption of power can be reduced.

4.3 DHT Based Service Discovery Protocol

If a node requests services, the node first searches its own local cache information in the following order: LSD, and NSC. If the corresponding services have been found from LSD, it means that the node itself is a service provider (SP) providing services. If those services have been searched in NSC, service information registered in NSC is used, which means that a service provider is physically neighbor node. If there exist

```
/* Find out the shortest node ranging from the current node to a target node
        Input: Looked up key (Key=Hashed Key), SD: Service Description
        LSD: Local Service Directory,
        NSC: Neighbor Service cache List,
*/
Search (Key, SD)
If   node has service in LSD
    reply service response message;
else If   one of the neighboring nodes in NSC has a service
         reply service response message;
       else forward Search(Key, SD) to logical node that takes charge of key;
```

Fig. 4. An algorithm for service discovery

two service providers or more in NSC, a service provider with the highest value in service power (slack time) is selected. If there are no registered services in local cache information, a service request message including the hash key value is created and sent to node in charge of such key value in DHT. Figure 4 shows an algorithm for service discovery.

The service response messages include information about IP address of a server providing those services. With regard to received service response message, a service provider with the highest value in service power (slack time) is selected so that the selected service provider can be requested for services to.

4.4 Cache Management

Several research works prove that a strong cache consistency is maintained in the single-hop-based wireless environment [12]. To maintain cache consistency, we use a simple weak consistency model based on a time-to-live mechanism, in which, if a routing node has not exceeded its service available time, it is considered to have a proper value. On other hand, if it does, it is eliminated from the table, because services are not effective any more. In this model, if a passing-by node has the same service name, the value is refreshed. If the cache size has not enough free space that a new service can not be added, service rarely referred to are deleted or eliminated in accordance with the LRU deletion policy.

5 Results and Analysis of Simulation

In this section, we evaluate the performance of mobile service protocol (MSD) proposed in this paper, comparing our protocol with those of the existing flooding-based service discovery and the DHT service discovery.

5.1 Environment of Simulation

To simulate service discovery in a mobile ad-hoc network environment, NS2 [20], a representative network simulation tool, is used. AODV protocol is adopted as a

Table 1. Simulation Parameters

Parameter	Value
Number of nodes	20 to 100
Network Area (x, y)	1500m x 1500m
Mean query generate time (secs)	5
Hello Message Interval (secs)	5
Routing Protocol	AODV
Movement average speed (m/s)	3 to 20
Movement maximum speed (m/s)	5 to 30
Mobility Pattern	Random way-point
Duration	300s
Transmission range	250m
Server Rate	30%
Mac	Mac/802_11

routing protocol, with 1500m x 1500m in the size of network area. The simulation is carried out with a change in the number of nodes ranging form 20 up to 100. A mobile scenario for mobile nodes where pause time on average is 2 seconds with average velocity and maximum velocity varying 3 to 20 m/sec and 5 to 30 m/sec, respectively, is designed for evaluating their performance. Table 1 shows the values of factors used in a simulation for algorithm of this section.

5.2 Results of Simulation

The average search hop count, a measurement standard that evaluates algorithm performance is the number of average hops needed for the search. More nodes cache more related information, locality and service response will get higher and quicker, respectively. When a service requester found out available services, the number of physical hops for paths is reduced, it can quickly find them out without going through many nodes and that power of node is relatively less consumed. Fig. 5 shows physical hop count from source node to destination node.

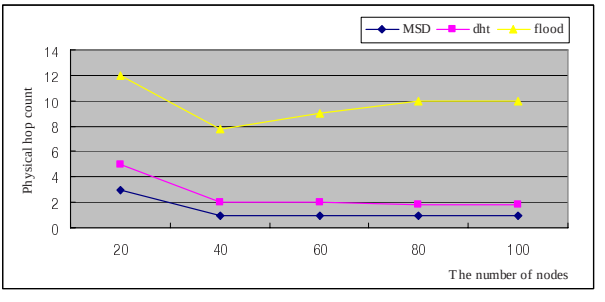


Fig. 5. Average physical hop count from source node to destination node

The search success ratio, a measurement standard that evaluates the performance and flexibility of an algorithm, is the percentage of the successful queries in the total queries requested. Fig. 6 shows success ratio in accordance with a change in the number of nodes if AODV is used as routing protocol and if, among those nodes, the ratio of service requesters is 10%. As shown in the figure, the approach we propose fared well on success ratio.

Fig. 7 shows the number of messages nodes communicate. In flooding method, a number of messages are communicated by broadcasting messages for service advertisements and service searches to neighboring nodes, which means that a transmission of messages is delayed with nodes consuming lots of power. DHT-based method finds out services from logical neighbors using log N method. But its application to a physical method would create a more number of messages.

In our proposed method, the number of messages communicated among nodes is a little more than DHT-based method, while the proposed method does not incur countless messages unlike flooding-based service discovery.

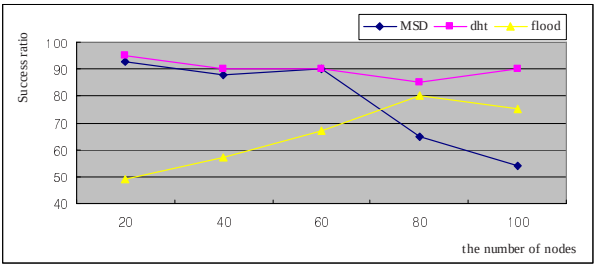


Fig. 6. Success ratio

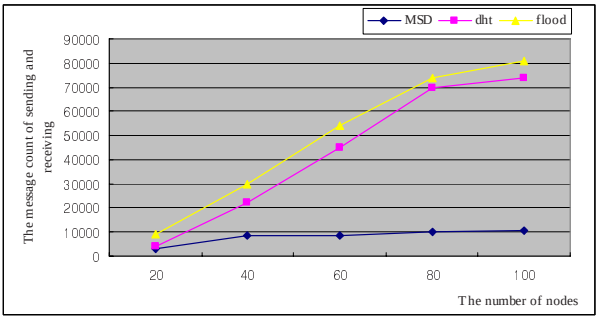


Fig. 7. Message count of sending and receiving

6 Conclusion

This paper proposes service discovery protocol in support of multi-hops using information of neighboring nodes based on DHT in an ad-hoc network. In the proposed service discovery method, the values of keys through hashing based on DHT method are created, and each node inserts a value “(Key, SD)” into a node in charge of the corresponding area. In addition to that, neighboring information, as about services, connection, power, location, etc., is shared through physically nearby neighboring nodes. The shared information finds out available services from its own nodes without communicating other nodes by using local cache information to an extent of more than 85% when a service requester would find them out.

A simulation has it that our proposed service discovery method is not affected by a velocity at which nodes move.

References

1. Toh, C.K.: Ad Hoc Mobile Wireless Networks: Protocols and Systems. Prentice Hall, Englewood Cliffs (2002)
2. The Gnutella web site, `http://www.gnutella.com`
3. The KaZaA web site, `http://www.kazaa.com`

4. Stoica, I., Morris, R., Karger, D., Kaashoek, M.F., Balakrishnan, H.: Chord: A scalable peer-to-peer lookup service for internet applications. In: SIGCOMM 2001: Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications, pp. 149–160. ACM Press, New York (2001)
5. Ratnasamy, S., Francis, P., Handley, M., Karp, R., Schenker, S.: A scalable content addressable network. In: SIGCOMM 2001: Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications, pp. 161–172. ACM Press, New York (2001)
6. Rowstron, A.I.T., Druschel, P.: Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems. Middle ware, 329–350 (2001)
7. Meier, R., Cahill, V., Nedos, A., Clarke, S.: Proximity-Based Service Discovery in Mobile Ad Hoc Networks. In: Kutvonen, L., Alonistioti, N. (eds.) DAIS 2005. LNCS, vol. 3543, pp. 115–129. Springer, Heidelberg (2005)
8. Guttman, E.: Service Location Protocol: Automatic Discovery of IP Network Services. IEEE Internet Computing 3, 71–80 (1999)
9. http://www.uddi.org
10. JiniArchitecturalOverview:TechnicalWhitePaper, http://www.sun.com/software/jini/whitepapers/architecture.html
11. Understanding UPnP: A White Paper. Microsoft Corporation (June 2000)
12. Cao, G.: Proactive Power-Aware Cache Management for Mobile Computing Systems. IEEE Trans. Computer 51(6), 608–621 (2002)
13. White paper.: Salutation Architecture: overview (1998), http://www.salutation.org/whitepaper/originalwp.pdf
14. Helal, S., Desai, N., Verma, V., Lee, C.: Konark-A Service Discovery and Delivery Protocol for Ad-hoc Networks. In: Proc. of the Third IEEE Conference on Wireless Communication Networks (WCNC), New Orleans (March 2003)
15. Helal, O., Chakraborty, D., Tolia, S., Kushraj, D., Kunjithapatham, A., Gupta, G., Joshi, A., Finin, T.: Allia:Alliance-based Service Discovery for Ad-Hoc Environments. In: Second ACM International Workshop on Mobile Commerce, in conjunction with Mobicom Atlanta GA, USA (2002)
16. Klein, M., Kognig-Ries, B., Obreiter, P.: Lanes - A lightweight overlay for service discovery in mobile ad hoc networks. In: 3rd Workshop on Applications and Services in Wireless Networks (2003)
17. Klemm, A., Lindermann, C., Waldhorst, O.P.: A Special-purpose peer-to-peer file sharing system for mobile ad hoc networks. IEEE VTC2003-Fall (2003)
18. Arau'jo, F., Rodrigues, L., Kaiser, J., Liu, C., Mitidieri, C.: CHR:a distributed hash table for wireless ad hoc networks. In: The 25th IEEE Int'l Conference on Distributed Computing Systems Workshops (DEBS 2002), Columbus, Ohio, USA (June 2005)
19. Perkins, C.E., Royer, E.M., Das, S.: Ad Hoc On-Demand Distance Vector Routing (AODV) Routing. RFC 3561, IETF (July 2003)
20. NS2 Object Hierarchy, http://www-sop.iniria.fr/planete/software/ns-doc/ns-current/aindex.html

Multivariate Option Pricing Using Quasi-interpolation Based on Radial Basis Functions*

Liquan Mei and Peipei Cheng

School of Science, Xi'an Jiaotong University
Xi'an, 710049, P.R. China

Abstract. Radial basis functions are well-known successful tools for interpolation and quasi-interpolation of the equal distance or scattered data in high dimensions. Furthermore, their truly mesh-free nature motivated researchers to use them to deal with partial differential equations(PDEs). With more than twenty-year development, radial basis functions have become a powerful and popular method in solving ordinary and partial differential equations now. In this paper, based on the idea of quasi-interpolation and radial basis functions approximation, a fast and accurate numerical method is developed for multi-dimensions Black-Scholes equation for valuation of european options prices on three underlying assets. The advantage of this method is that it does not require solving a resultant full matrix, therefore as indicated in the the numerical computation, this method is effective for option pricing problem.

Keywords: radial basis functions, options pricing, quasi-interpolation, DSMQ.

Classification: AMS(2000) 90A09, 65M12.

1 Introduction

The valuation and hedging of financial option contracts is a subject of considerable practical significance. The holders of such contracts have the right to undertake certain actions so as to receive certain payoffs. The valuation problem consists of determining a fair price to charge for granting these rights. A related issue, perhaps of even more importance to practitioners, is how to hedge the risk exposures which arise from selling these contracts. An important feature of such contracts is the time when contract holders can exercise their right. If this occurs only at the maturity date of the contract, the option is classified as "European". If holders can exercise any time up to and including the maturity data, the option is said to be "American". The value of a European option is given by the solution of the Black-Scholes PDE[1].

In current pratice, the most common method of solving option pricing problems, including radial basis functions method(more details in [4]). In recent years,

* The project is supported by NSF of China(10471109).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 620–627, 2008.
© Springer-Verlag Berlin Heidelberg 2008

radial basis functions have become a powerful and popular tool in solving ordinary and partial differential equations. Numerous interesting numerical results have been obtained for solving PDEs appeared in different application areas, such as the biphasic mixture model for tissue engineering problems, heat transfer, nonlinear Burgers' equation, shallow water equation for tide and current simulation, etc. Some theoretical results are also obtained for RBFs meshless methods. Wendland [2] proposed meshless Galerkin methods using RBFs for solving second order elliptic problem, and proved the convergence and the same error bounds in the energy norm as the classical finite element methods. Recently, with the development of China financial market, valuation of options is worthwhile to do further research. In this paper, we extend the multilevel univariate quasi-interpolation formula proposed in [3] to multidimensions using the dimension-splitting multiquadric (DSMQ) basis function approach to solve the Black-Schloes (B-S) equation.

The organization of this paper is as follows. In section 2, we introduce the B-S equation and radial basis functions. In the following section 3, the numerical solution with the DSMQ method is given. Lastly, section 4 is our conclusion.

2 Black-Scholes Equation and Two Interpolation Formulas

In 1973, Black-Schloes proposed an explicit formula for evaluating European call options without dividends. By assuming that the asset price is risk-neutral, Black and Scholes showed that the European call options value satisfies a log-normal diffusion type partial differential equation which is known now as the Black-Scholes equation.

We consider the following three dimensional Black-Scholes equation

$$\frac{\partial C}{\partial t} = -\sum_{i=1}^{3} r s_i \frac{\partial C}{\partial s_i} - \frac{1}{2}\sum_{i,j=1}^{3} \sigma_i \sigma_j \rho_{ij} \frac{\partial^2 C}{\partial s_i \partial s_j} + rC, \tag{1}$$

where r is the risky-free interest rate, $\sigma_1, \sigma_2, \sigma_3$ is the volatility of the price s_1, s_2, s_3, and $C(s_1, s_2, s_3, t)$ is the option value at time t and stock price s_1, s_2, s_3. The terminal condition is special for each type of payoff. For example, in the case of a European call on the maximum of three assets with strike price E, maturity T, the condition is given by

$$C(s_1, s_2, s_3, T) = (\max\{s_1(T), s_2(T), s_3(T)\} - E)_+.$$

Before discretizing (1), the PDE should be simplified. Introducing the changes of variables $x = \log(s_1/E), y = \log(s_2/E)$, $z = \log(s_3/E), \tau = T - t$, and $C(s_1, s_2, s_3, t) = Eu(x, y, z, \tau)$ leads to the forward parabolic equation with constant coefficients

$$\begin{aligned}\frac{\partial u}{\partial \tau} &= (r - \frac{\sigma_x^2}{2})\frac{\partial u}{\partial x} + (r - \frac{\sigma_y^2}{2})\frac{\partial u}{\partial y} + (r - \frac{\sigma_z^2}{2})\frac{\partial u}{\partial z} + \frac{\sigma_x^2}{2}\frac{\partial^2 u}{\partial x^2} + \frac{\sigma_y^2}{2}\frac{\partial^2 u}{\partial y^2} \\ &+ \frac{\sigma_z^2}{2}\frac{\partial^2 u}{\partial z^2} + \sigma_x\sigma_y\rho_{xy}\frac{\partial^2 u}{\partial x \partial y} + \sigma_x\sigma_z\rho_{xz}\frac{\partial^2 u}{\partial x \partial z} + \sigma_y\sigma_z\rho_{yz}\frac{\partial^2 u}{\partial y \partial z} - ru,\end{aligned} \tag{2}$$

with initial condition

$$u(x, y, z, 0) = (\max\{e^x - 1, e^y - 1, e^z - 1\})_+.$$

One of most popular methods solving equation (2) is the finite difference method (see [5]), but we adopt another effective method named quasi-interpolation with RBFs in this paper. There are many RBFs (expression of ϕ) available. The most commonly used are

MQs : $\phi(r) = (r^2 + c^2)^{\frac{2k-1}{2}}, \quad k \in N,\ c > 0,$

Thin plate-splines : $\phi(r) = r^2 \log(r),$

Gaussians : $\phi(r) = e^{-cr^2}, \quad c > 0,$

Inverse MQs : $\phi(r) = (r^2 + c^2)^{-\frac{2k-1}{2}}, \quad k \in N,\ c > 0,$

where $r = \|x - x_i\|_2, c$ is a constant.

Among the four RBFs, MQ, which was first presented by Hardy, is used extensively. Franke did a comprehensive study on various RBSs, and found that MQ generally performs better for the interpolation of 2D scattered data. Therefore, in this work, we will concentrate on MQ RBFs.

Definition 1. *A function $F : \Re^n \to \Re$ is conditional positive definite of order k if for all finite subsets Ξ from $\Re^n$, the quadratic form*

$$\sum_{\xi,\zeta} \lambda_\zeta \lambda_\xi F_{\xi\zeta}, \quad \xi, \zeta \in \Xi \tag{3}$$

is nonnegative for all $\lambda = \{\lambda_\xi\}_{\xi\in\Xi}$ which satisfy $\sum_{\xi\in\Xi} \lambda_\xi q(\xi) = 0$ for all $q \in P_n^{k-1}$. F is strictly conditional positive definite of order k if the quadratic form (3) *is positive for all nonzero vector λ.*

If we are supplied with a finite set of interpolation points $X \subset \Re^n$ and a function $f : X \to \Re$, we can construct an interpolant to f of form

$$(Sf)(x) = \sum_{x_i \in X} \lambda_i \phi(\|x - x_i\|) + p(x), \quad x \in \Re^n,\ p \in P_n^{k-1}.$$

Of course, for Sf , the real number λ_i must be chosen to satisfy the system

$$\begin{cases} (Sf)(x_i) = f(x_i) & \text{for} \quad x_i \in X, \\ \sum\limits_{x_i \in X} \lambda_i x_i^\alpha = 0 & \text{for} \quad |\ \alpha\ | < k. \end{cases} \tag{4}$$

Here, the purpose we add polynomial p is to satisfy the nonsingularity of the interpolation matrix. We know that we have a unique interpolant Sf of f if $\phi(r)$ is a conditional positive definite radial basis function of order k. The side-condition

in (4) is used to take up the extra degrees of freedom that are introduced through the use of polynomial p.

According to the process of interpolation, we adopt a special MQ function $\phi(r) = \sqrt{r^2 + c^2}$ in our paper, whose nonsingularity can be deduced from the following theory.

Theorem 1. *Let $g \in C^{\infty}[0, \infty]$ be satisfied that g' is completely monotonic but not constant. Suppose further that $g(0) \geq 0$. Then the interpolation matrix is nonsingular for $\phi(r) = g(r^2)$.*

The result is due to Micchelli (1986), which can be applied to the MQ, namely $g(r) = \sqrt{r + c^2}$ and $\phi(r) = \sqrt{r^2 + c^2}$, then it gives the desired nonsingular result [4]. Thus, for Gaussians, inverse MQs whose interpolation matrices are positive definite and $\phi(r) = \sqrt{r^2 + c^2}$ we can adopt the simpler form for convenience.

$$(Sf)(x) = \sum_{x_i \in X} \lambda_i \phi(\|x - x_i\|), \quad \text{for} \quad x \in \Re^n.$$

Now let's introduce the quasi-interpolation formula in the ground of the basic knowledge of the interpolation above.

Given the data $\{x_i, F_i\}_{i=0}^n$, $x_i \in X \subset \Re^n$, we have

$$(Sf)(x) = \sum_{x_i \in X} F_i \psi(\|x - x_i\|), \text{for } x \in \Re^n,$$

where ψ is the linear combination of the RBFs or adding polynomials sometimes. Compared with (4), it need not satisfy the interpolation condition, and the coefficients F_i are known from the given data. Therefore, it is no need of solving linear equations to get the coefficients, which not only reduces the computation but also avoids the consideration of the nonsingularity of the interpolation matrix. Furthermore, it achieves essentially the same approximation orders as the interpolation in the set of girded data. Then it is almost as attractive as the interpolation in many cases, becoming more and more popular in nowadays researches.

3 Solve European Style Options Pricing Equation on Three Assets

Now we describe the numerical solution for the PDE (2) which requires the solution of a large block triangular linear system.

Given data $\{x_i, F_i\}_{i=0}^n, x_0 \leq x_1 \leq \ldots \leq x_n$, Wu-Schaback's formula is given by

$$\Im F(x) = \sum_{i=0}^{n} F_i \alpha_j(x), \tag{5}$$

where the interpolation kernel $\alpha_j(x)$ is formed from linear combination of the MQ basis functions, plus a constant and linear polynomial.

$$\begin{aligned}
\alpha_0(x) &= \frac{1}{2} + \frac{\phi_1(x) - (x - x_0)}{2(x_1 - x_0)}, \\
\alpha_1(x) &= \frac{\phi_2(x) - \phi_1(x)}{2(x_2 - x_1)} - \frac{\phi_1(x) - (x - x_0)}{2(x_1 - x_0)}, \\
\alpha_{n-1}(x) &= \frac{x_n - x - \phi_{n-1}(x)}{2(x_n - x_{n-1})} - \frac{\phi_{n-1}(x) - \phi_{n-2}(x)}{2(x_{n-1} - x_{n-2})}, \\
\alpha_n(x) &= \frac{1}{2} + \frac{\phi_{n-1}(x) - (x_n - x)}{2(x_n - x_{n-1})}, \\
\alpha_j(x) &= \frac{\phi_{j+1}(x) - \phi_j(x)}{2(x_{j+1} - x_j)} - \frac{\phi_j(x) - \phi_{j-1}(x)}{2(x_j - x_{j-1})}, \quad j = 2, \ldots, n-2.
\end{aligned} \tag{6}$$

It is shown that (5) preserves monotonicity and convexity. Furthermore, Wu and Schaback concluded that no improvement towards $O(h^2)$ convergence is possible just by the changing of end condition or data points placement.

We consider the function $u(x, y, z, \tau)$ at the discrete set of points

$$\begin{aligned}
x_i &= L_x + i\triangle_x, \quad i = 0, 1 \ldots, N_x, \quad y_j = L_y + j\triangle_y, \quad j = 0, 1 \ldots, N_y, \\
z_k &= L_z + k\triangle_z, \quad k = 0, 1 \ldots, N_z, \quad \tau_m = (m-1)\triangle_\tau, \quad m = 1, \ldots, M,
\end{aligned}$$

where L_x, L_y and L_z are the lower bounds of the spatial grid, $\triangle_x, \triangle_y$ and $\triangle_z$ define the spatial grid steps, $\triangle_\tau$ is the time step, N_x, N_y, N_z and M are the number of spatial and time steps. Similar to formula (6), we can apply $\alpha_j(x)$ to variable y, z , getting

$$\begin{aligned}
\beta_j(y) &= \alpha_j(y), \quad j = 0, 1 \ldots, N_y, \\
\gamma_k(z) &= \alpha_k(z), \quad k = 0, 1 \ldots, N_z.
\end{aligned}$$

we extend the Wu and Schaback univariate quasi-interpolation formula to three dimensions on rectangular grid. Our three dimensional qusi-interpolation scheme uses the dimension-splitting multiquadric (DSMQ) basis functions.

Given the data $(x_i, y_j, z_k, F_{ijk}(\tau))$, the three dimensional quasi-interpolation formula is given by

$$\Im F(x, y, z, \tau) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} \sum_{k=0}^{N_z} F_{ijk}(\tau)\alpha_i(x)\beta_j(y)\gamma_k(z), \tag{7}$$

where $\alpha_i(x)$ is defined by (6).

Collocating equation (2) at the $(N_x+1)*(N_y+1)*(N_z+1)$ points (x_i, y_j, z_k),$i = 0, 1 \ldots, N_x, j = 0, 1 \ldots, N_y, k = 0, 1 \ldots, N_z$, and bearing in mind the approximation (7), the following system of linear equations (8) is obtained.

$$\frac{\partial u(x_i, y_j, z_k, \tau)}{\partial \tau} = (r - \frac{\sigma_x^2}{2})\frac{\partial u(x_i, y_j, z_k, \tau)}{\partial x} + (r - \frac{\sigma_y^2}{2})\frac{\partial u(x_i, y_j, z_k, \tau)}{\partial y}$$

$$
\begin{aligned}
&+(r-\frac{\sigma_z^2}{2})\frac{\partial u(x_i,y_j,z_k,\tau)}{\partial z}+\frac{\sigma_x^2}{2}\frac{\partial^2 u(x_i,y_j,z_k,\tau)}{\partial x^2}\\
&+\frac{\sigma_y^2}{2}\frac{\partial^2 u(x_i,y_j,z_k,\tau)}{\partial y^2}+\frac{\sigma_z^2}{2}\frac{\partial^2 u(x_i,y_j,z_k,\tau)}{\partial z^2}\\
&+\sigma_x\sigma_y\rho_{xy}\frac{\partial^2 u(x_i,y_j,z_k,\tau)}{\partial x\partial y}+\sigma_x\sigma_z\rho_{xz}\frac{\partial^2 u(x_i,y_j,z_k,\tau)}{\partial x\partial z}\\
&+\sigma_y\sigma_z\rho_{yz}\frac{\partial^2 u(x_i,y_j,z_k,\tau)}{\partial y\partial z}-ru(x_i,y_j,z_k,\tau).
\end{aligned}
\tag{8}
$$

Since the basis functions do not depend on time, the time derivative of u is simply the time derivative of the coefficients

$$
\frac{\partial u(x_i,y_j,z_k,\tau)}{\partial \tau}=\sum_{i=0}^{N_x}\sum_{j=0}^{N_y}\sum_{k=0}^{N_z}\frac{u_{ijk}(\tau)}{\partial \tau}\alpha_i(x)\beta_j(y)\gamma_k(z).
$$

The first and second partial derivatives of u with respect to x are given by

$$
\frac{\partial u(x_i,y_j,z_k,\tau)}{\partial x}=\sum_{i=0}^{N_x}\sum_{j=0}^{N_y}\sum_{k=0}^{N_z}u_{ijk}(\tau)\frac{\partial \alpha_i(x)}{\partial x}\beta_j(y)\gamma_k(z),
$$

$$
\frac{\partial^2 u(x_i,y_j,z_k,\tau)}{\partial x^2}=\sum_{i=0}^{N_x}\sum_{j=0}^{N_y}\sum_{k=0}^{N_z}u_{ijk}(\tau)\frac{\partial^2 \alpha_i(x)}{\partial x^2}\beta_j(y)\gamma_k(z).
$$

In the same way, partial derivatives of u with respect to y, z can be also computed from $\beta_j(y), \gamma_k(z)$. In the matrix form, equations (8) can be expressed as

$$
\begin{aligned}
\Psi\dot{U}=&(r-\frac{\sigma_x^2}{2})\Psi_x U+(r-\frac{\sigma_y^2}{2})\Psi_y U+(r-\frac{\sigma_z^2}{2})\Psi_z U+\frac{\sigma_x^2}{2}\Psi_{xx}U+\frac{\sigma_y^2}{2}\Psi_{yy}U\\
&+\frac{\sigma_z^2}{2}\Psi_{zz}U+\sigma_x\sigma_y\rho_{xy}\Psi_{xy}U+\sigma_x\sigma_z\rho_{xz}\Psi_{xz}U+\sigma_y\sigma_z\rho_{yz}\Psi_{yz}U-r\Psi U,
\end{aligned}
\tag{9}
$$

where U denotes the vector containing the unknown option value

$$
u_m=u(x_i,y_j,z_k,\tau),\qquad m=i*N_y*N_z+j*N_z+k
$$

and $\Psi, \Psi_x, \Psi_y, \Psi_z, \Psi_{xx}, \Psi_{yy}, \Psi_{zz}, \Psi_{xy}, \Psi_{xz}, \Psi_{yz}$ are the $(N_x+1)*(N_y+1)*(N_z+1)$ order matrices.

For fixed points (x_i, y_j, z_k), equation (9) is a linear system of first-order homogeneous ordinary differential equations with constant coefficients. Starting from the initial condition, we can use any backward time integration scheme to obtain the unknown coefficients U at each time step $n\triangle\tau$. For notational convenience, let U_n denotes the vector$[(u_{ijk}^{n\triangle\tau})_{N_x,N_y,N_z}]^T$ at each time step, and

$$
\begin{aligned}
P=&(r-\frac{\sigma_x^2}{2})\Psi_x+(r-\frac{\sigma_y^2}{2})\Psi_y+(r-\frac{\sigma_z^2}{2})\Psi_z+\frac{\sigma_x^2}{2}\Psi_{xx}+\frac{\sigma_y^2}{2}\Psi_{yy}+\frac{\sigma_z^2}{2}\Psi_{zz}\\
&+\sigma_x\sigma_y\rho_{xy}\Psi_{xy}+\sigma_x\sigma_z\rho_{xz}\Psi_{xz}+\sigma_y\sigma_z\rho_{yz}\Psi_{yz}-r\Psi.
\end{aligned}
$$

The following implicit numerical time integration scheme is used to discretize equation (9) for the valuation of the European options

$$\Psi U_n = \Psi U_{n-1} + \triangle\tau P[\theta U_{n-1} + (1-\theta)U_n]. \tag{10}$$

Let

$$P_1 = \Psi - (1-\theta)\triangle\tau P, \quad P_2 = \Psi + \theta\triangle\tau P.$$

Equation (10) can be rewritten as

$$P_1 U_n = P_2 U_{n-1}.$$

Here, to deal with the ill-condition of the linear system, we have

$$P_{1\lambda} = (\lambda P_1 - P_2)^{-1} * P_1, \quad P_{2\lambda} = (\lambda P_1 - P_2)^{-1} * P_2,$$

where λ is the constant that satisfies $\lambda P_1 - P_2$ is nonsingular. In the following computation, we choose θ in equation (10) is 0.5, and $r = 0.1, T = 1$.

Table 1. Pricing results for different volatility on three assets

σ_x	σ_y	σ_z	ρ_{xy}	ρ_{xz}	ρ_{yz}	RBFs	Johnson	error
0.30	0.30	0.30	0.90	0.90	0.90	15.319	15.436	0.00117
0.30	0.30	0.30	0.60	0.40	0.60	18.228	18.245	0.00023
0.25	0.30	0.35	0.90	0.90	0.90	9.235	9.223	0.00012
0.25	0.30	0.35	0.60	0.40	0.60	11.969	11.929	0.0004

Table 1 shows the numerical results for different volatilities and correlation coefficients of quasi-interpolation using RBFs and that of given by Johnson [8] towards B-S equation. With the comparison we know the RBFs method is very effective to solve PDEs. Table 2 gives the numerical option pricing results for different asset prices respectively.

Table 2. Pricing results for different asset prices

Ps_1	Ps_2	Ps_3	E	RBFs	Johnson	error
40	40	40	35	12.321	12.384	0.00063
40	40	40	45	6.252	6.270	0.00018
40	45	50	35	19.430	19.448	0.00088
40	45	50	45	11.821	11.882	0.00061

In Table 1, the second and third rows are in condition $s_1 = 40, s_2 = 45, s_3 = 50, E = 40$. The latter two rows are in condition $s_1 = 40, s_2 = 40, s_3 = 40, E = 40$. In Table 2, $\sigma_x = \sigma_y = \sigma_z = 0.3,\ \rho_{xy} = \rho_{xz} = \rho_{yz} = 0.9$.

4 Conclusion

Based on the idea of quasi-interpolation and radial basis functions approximation, a fast and accurate numerical method is developed for multidimensions Black-Scholes equation for valuation of european options prices on three underlying assets. Since this method does not require solving a resultant full matrix, the ill-conditioning problem resulting from using the radial basis functions as a global interpolant can be avoided. The method has been shown to be effective in solving option pricing problem.

The positive constant c^2 contained in the MQ function is called a shape parameter whose magnitude of value affects the accuracy of the approximation. In most applications of using the MQs for scattered data interpolation or quasi-interpolation, a constant shape is assumed for simplicity. It has shown that the use of the MQs for solving partial differential equations is highly effective and accurate. However, the accuracy of the MQs is greatly affected by the choice of a shape parameter whose optimal value is still unknown. In this paper, we adopt value $c^2 = 4h^2$, in which h describes the density of the scattered data. Towards the importance and the difficult choice of the shape parameter in MQs method, we will focus more attention on the analysis of the shape parameter in the future work.

References

1. Wilmott, P., Dewynne, J., Howison, S.: Option Pricing: mathematical models and computation. Oxford Financial Press (1993)
2. Wendland, H.: Meshless Galerkin method using radial basis functions. Mathematics of Computation 68(228), 1521–1531 (1999)
3. Ling, L.: Multivariate quasi-interpolation schemes for dimension-splitting multiquadric. Applied Mathematics and Computation 161, 195–209 (2005)
4. Buhmann, M.: Radial Basis Functions. Cambridge University Press, Cambridge (2003)
5. Mei, L., Li, R., Li, Z.: Numerical solutons of partial differential equations for trivariate option pricing. J. of Xi'an Jiaotong University 40(4), 484–487 (2006)
6. Golub, G., Van Loan, C.: Matrix Computation, pp. 218–225. The Johns Hopkins University Press, Baltimore (1996)
7. Kansa, E.: Multiquatrics-A scattered data approximation scheme with applications to computational fluid-dynamics-H. Solutions to parabolic, elliptic and hyperbolic partial differential equations. Computers and Mathematics with Applications 19(6-8), 147–161 (1991)
8. Johnson, H.: Options on the Maximum or Minimum of Several Assets. Journal of Financial and Quantitative Analysis, 227–283 (1987)

A Novel Embedded Intelligent In-Vehicle Transportation Monitoring System Based on i.MX21

Kaihua Xu[1], Mi Chen[1], and Yuhua Liu[2]

[1] College of Physical Science and Technology, Huazhong Normal University, Wuhan, 430079, China
[2] Department of Computer Science, Huazhong Normal University, Wuhan, 430079, China
chenmi0201@mails.ccnu.edu.cn

Abstract. In this paper, we introduces a novel intelligent vehicle transportation monitoring system (NIVTS) based on i.MX21 designed and realized by us, it overcomes drawbacks of traditional vehicle monitoring and it has characteristics of multi-functions, real time. It uses many technologies fusing GPS, GPRS, WebGIS, RS, and is composed mainly by two parts: the Mobile Terminal and Monitoring Center. Especially, GPRS technology is the reliable support for wireless communication smooth transition; then, the principles and the key technologies of Monitoring System are discussed respectively in detail. A novel vehicle scheduling model based on ant colony optimization algorithms are proposed and the course of scheduling tasks is also introduced. The system not only can be used in the vehicles scheduling, but also in the vessel and other fields, and it will reduce workloads of traffic control operators. The solid technical support will be provided for the ITS research.

1 Introduction

Intelligence transportation System (ITS) plays a critical role in the development of social economy and modern city climbing. The highway traffic system also changes more and more complex, faced with annually increasing demand for travel and transport of goods, the transportation system is reaching the limits of existing capacity [1]. ITS can help ease this strain, and reduce the emissions created and fuel wasted in associated congestion, through the application of modern information technology and communications. In our country, the search on ITS started late but it has entered a fast-developing stage in recent years. ITS is gradually being applied to transport companies, freight transportation, and many travel service business as researches going deep. In this article, an Embedded Intelligent Vehicle Monitoring and Management System is a service ITS dealing with logistic problems of the vehicles, which is the reliable technical support for wireless communication smooth transition.

2 Structure and Functions of NIVTS

In this paper, NIVTS uses many technologies fusing GPS, GPRS, GIS, RS, and is composed by two major parts: the Mobile Terminal and the Monitoring Center. While

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 628–635, 2008.
© Springer-Verlag Berlin Heidelberg 2008

passing on the data from the Mobile Terminal to Monitoring Center, the vehicles positional information and image data, which are gained separately through the GPS and sensor, are processed firstly by the control unit of the Mobile Terminal, and then compressed, packed, passed to the GPRS communication module and transmitted to GPRS wireless Internet with the GPRS antenna, and next passed on in to China Mobile GPRS platform automatically by wireless Internet, finally carried between the platform and the Monitoring Center through the wired Internet. It is the counter process of passing on the data from the Monitoring Center to the Mobile Terminal, and the system architecture is as shown in Fig 1. Although the system architecture is quite complex, in practice research, except for the ready-made network and service platform, the system may be considered as two major parts: the vehicle Mobile Terminal and the Monitoring Center [2] .

NIVTS has implemented many functions, especially mainly include “Real-time monitoring and management of overspend, overfreight and overtime driving”, The information can be relayed via the World Wide Web to any desktop PC with Internet access where it can be viewed by the user, by using a secure password system . The Vehicle Monitoring system can show all kinds of information, including the position of vehicles at any time, the time taken to complete journeys and how long the vehicle spent at a particular location. Vehicle Monitoring brings you a well designed, efficiency, cost effective vehicle management solution.

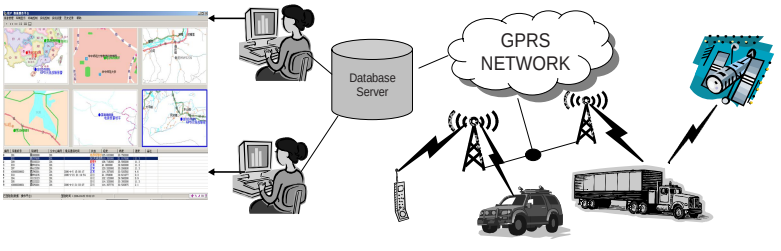


Fig. 1. The architecture of vehicle monitoring system

3 The Design and Realization of Vehicle Mobile Terminal

The hardware design and the software writing are included in Mobile Terminal.

3.1 The Component of the Mobile Terminal Hardware

Many functional modules are mainly contained within terminal hardware (Fig. 2): namely the CPU module (MC9328MX21), the GPRS module, the GPS module, the power management unit (PWU), the peripheral NanFlash, SDRAM, the CMOS Sensor, the TFT LCD, the polymer lithium-Lon, the multiplex A/D conversion and the vehicle telephone [7]. MC9328MX21 (i.MX21) provides a leap in performance with an ARM926EJ-S microprocessor core that provides native security and accelerated Java support in addition to highly integrated system functions. The i.MX21 processor features the advanced and power-efficient ARM926EJ-S core operating at speeds up

to 266 MHz, and comes with an enhanced Multimedia Accelerator (eMMA) comprising an IEC14496-2 compliant MPEG-4 encoder and decoder, independent preprocessing and post-processing stages which provide exceptional image and video quality.

Among them, the stable and continual electric power is provided by the power management unit and polymer lithium-Lon battery for the entire hardware system to guarantee system energy power; the system central module is the CPU that not only regulates the other modules to work effectively, but also coordinate the normal communication among system interior modules; the important function unit is the GPS module to calculates its position data [6], and the realization of communication between the GPRS network and the Mobile Terminal is the GPRS module; Besides what is said above, it also contains the functions of RS image data gain, self-monitoring and so on. That has met the need of Intelligence vehicle scheduling nowadays.

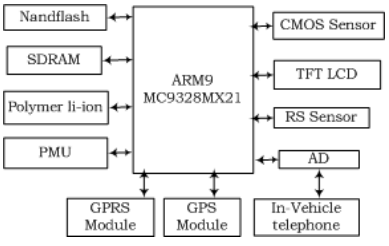


Fig. 2. The structure of vehicle Mobile Terminal

3.2 The Principle of Vehicle Mobile Terminal

The system adopts WINCE5.0 Embedded OS, porting Wince Kernel and various drivers(UART、LCD、IIS、RTC、Codec、SD, etc.) by Platform Builder , complier the application program by EVC[4]. The BSP develop flowcharts is bootloader (Uboot), Kernel (Wince),drivers ,application code(see Fig. 3) The functions of Mobile Terminal are mainly realized by the GPS module, the GPRS module and the other modules all together. The vehicles positional information and image data inside and outside, which are gained separately through the GPS antenna and the sensor, are processed firstly by the control unit of the Mobile Terminal, then compressed, packed, passed to the GPRS module and transmitted to GPRS wireless Internet with the GPRS antenna of GPRS communication module. Among them, the real-time dynamic latitude of the Mobile Terminal, using globe's real-time position data which is received from the GPS satellite through GPS antenna of the terminal and regarded as localization reference system and the information source, are figured out by the GPS module which has already placed in it. The results are transmitted to the control unit of the terminal; moreover, not only may each instruction and the information which are sent from the Monitoring Center be done, but also can the operation request be passed upwards on to the Monitoring Center after driver and passengers pressing the buttons by the control unit of Mobile Terminal. Thus the scheduling, management and monitoring for the Mobile Terminal are realized by the Monitoring Center.

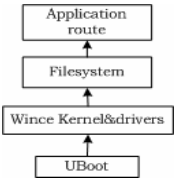


Fig. 3. BSP Block Diagram of i.MX21

4 The Design and Realization of Monitoring Center

The main purpose of Monitoring Center design is configuring the software based on building hardware platform successfully. Its central software includes: LAN software of Monitoring Center, Internet communication software, map database software, BS workstation service software and the others. Normal communication and reasonable call among the main central software are guaranteed by the LAN software of Monitoring Center.

4.1 The Structure of Monitoring Center

The monitoring center hardware platform contains communication server, database server, GIS/RS workstations and the other peripheral equipments, and other network communication hardware is realized through the LAN which is set up by the Monitoring Center. The work relation of the monitoring center software (see Fig. 4) which is installed in the corresponding hardware equipment is cascading: communication relevant information is received and retransmitted by Internet communication software; the vehicles dynamic data, the image data and the others are stored in the database function software. Relevant information of the geography data is in GIS/RS workstation service software which is used to process the long-distance position data and so on.

4.2 The Principle of Monitoring Center

The functions of Monitoring Center are mainly implemented by communication server, the database server, GIS/RS workstations and the other peripheral equipments together. The Mobile Terminal messages which are passed on with Internet are retransmit to the database server and GIS/RS workstations according to category, at the same time, the information and the instruction which are sent from GIS/RS workstation to the Terminal are transmitted by communication server that is connected with Internet the of Monitoring Center . All kinds of vehicles dynamic real-time information which are sent with server are contained in the database server, that are the GPS positional information, pictures, the vehicles condition information, the keyboard information, the map data, the vehicles file data, the user management data and documents and so on. The control platforms and unit of the operating are GIS/RS workstations. The operator on the one hand may judge from the displayed information above and send the communication service instructions to operate the vehicles, on the

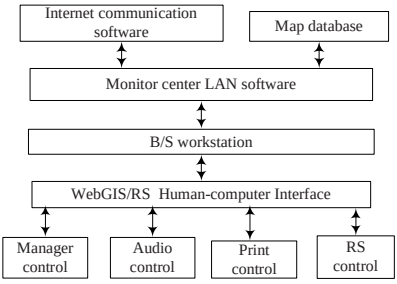


Fig. 4. The logic structure of Monitoring Center

other hand may command the operating control platform according to the demands of managers and users, and the instructions are implemented in Mobile Terminal through communication server.

5 The Intelligent Scheduling Algorithm of NIVTS

5.1 The Review of the Ant Algorithm

Observing the searching food process of ant colony, entomologist found that ant could release a sort of special secretion in the move track, the behind ants can choose the way of it based on the secretion left by the before ones. The more secretion on one way, the bigger probability the ants will choose that way. So, the collectivity actions of the ant colony form a sort of positive feedback mechanism. Through such information interchange and cooperating between one and another to find the shortest way to the food [3-4].

Because vehicle scheduling is the same as ants looking for food, and has the same characteristic at self-organization economic activities which take the community profit will be produced in this disorderly process, it is one of intelligent scheduling methods using the ant algorithms to solve the vehicles scheduling problem.

5.2 The ACS Model of Transport Road Network

A transport road can be represented by a digraph $G=(V,E,W)$, where V represents the set of vertices, E represents the set of edges and W represents the set of weight. First, we give a hypothesis to the vehicle source and destination to the graph node i, j, and have M road supply to choice represents l_{ij} (Fig. 5). η_{ij}^{k} represented part visibility, namely the attract degree to vehicle k of road l_{ij}; $\tau_{ij}(t)$ represents secretion pheromone. And, we proposed $\tau_{ij}(0)$ is equal to a constant ε. As time going, the secretion pheromone in a road decrease, we use parameter ρ represented the volatilization rate of the quantity of pheromone on road l_{ij}.

So, $\tau_{ij}(t)$ will fresh at N time slice as the following formula (1).

$$\tau_{ij}(t+N) = (1-\rho)\cdot\tau_{ij}(t) + \Delta\tau_{ij} \,. \tag{1}$$

The pheromone that vehicle k left on road l_{ij} from the time t to time $t+N$ represented by formula (2).

$$\Delta\tau_{ij} = \sum_{k=1}^{N} \Delta\tau_{ij}^{k} \,. \tag{2}$$

where $\Delta\tau_{ij}^{k}$ (as the following formula 3) represents the quantity of pheromone laid on certain road l_{ij} by scheduling vehicle k.

$$\Delta\tau_{ij}^{k} = \begin{cases} \dfrac{Z}{C_k(t)} & \text{if vehicle } k \text{ choose road } l_{ij} \\ 0 & \text{otherwise} \end{cases} \,. \tag{3}$$

where Z is a constant and $C_k(t)$ is the cost of vehicle k at the time t.

At the time t, the probability that vehicle k choose the road l_{ij} is $P_{ij}^{k}(t)$ as the following formula (4). That considers two elicitation gene α, β, one is α represented the strength of pheromone, the other one is β, represented part visibility, namely the attract degree to ant.

$$P_{ij}^{k}(t) = \frac{\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}^{k}\right]^{\beta}}{\sum_{j\in\{1,2,\ldots,M\}}\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}^{k}\right]^{\beta}} \tag{4}$$

We get the probability of every road with formula (4), the find the optimization road using every road probability.

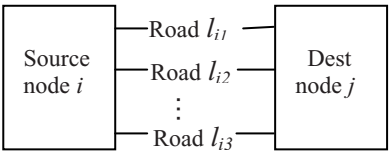


Fig. 5. The Road model of Ant Algorithm

6 The Solution of a Freight Scheduling Task

Give a hypothesis that *N* freights should be transported and *M* vehicles can be scheduled, and the *N* freights are divided into *N* steps with only one freight being done in the every. If it is hoped that the cost of the whole task is lowest, every step should be cheapest. Following variables are introduced for modeling.

(1) The distance d_{ij} of finishing the freight j transported by vehicle i;

(2) The time λ_{ij} of freight j transported by vehicle i;

(3) The waste h_{ij} of freight j loaded by vehicle i;

(4) The tough degree s_{ij} of path when freight j is transported by vehicle i;

(5) The conditions w_i of vehicle i

The cost of freight i transported by vehicle i is c_{ij}, with

$$c_{ij} = a_1 h_{ij} + a_2 d_{ij} + a_3 s_{ij} + a_4 \lambda_{ij} + a_5 w_i \ . \tag{5}$$

While a_1, a_2, a_3, a_4, a_5 are constants. Given the whole transport cost is C, if it is hoped to be lowest, the cost model of the intelligent scheduling is built (as follows formula 6), and this formula is other represent way of $C_k(t)$ in the front formula 3

$$\mathrm{C} = \min \sum_{j=1}^{N} c_{ij} \ . \text{ and } i \in \{1, 2, 3, \cdots, \ \mathrm{M}\} \tag{6}$$

A scheduling task can be divided into 4 steps according to the study above in practice.

Step 1. The primary information is obtained through users, enterprises and public information networks, company marketing network.

Step 2. Input geographic names and the user data in the database server transferred, the function of the group call and scheduling being locked in the certain scope willfully is used to search and transfer all vehicles data of around goods in the surrounding area of origin of goods.

Step 3. Input the data of goods and transports resources, the best scheduling plan is calculated based on ant algorithms intelligent scheduling software. The primary information and the scheduling instructions are sent to the chosen vehicle by using the function of mutually passing permanent real-time on-line information.

Step 4. Using the function module of "Transmission of real-time images on the whole trip", "Real-time playback of data about vehicles position, speed, time and state on the whole trip", "Real-time records of the whole trip about the vehicles passing on the data" and so on, freight transportation are automatic tracked and recorded on whole trip. Then, all data of the on-duty vehicles are saved in resources class of database, which is prepared for the next round scheduling plan.

7 Conclusions

In this paper, many technologies such as GPS, GPRS, GIS and RS are used, and the alarming and many other functions are included in NIVTS with ARM processor i.MX21. The vehicle utilization will be effectively enhanced and the transport cost will be reduced to a certain extent, so that the vehicles of Transport Company are easy to be managed and scheduled. In addition, because ant algorithms are used in the system, a good effect is achieved in the course of practical operation. The system not only can be used in the vehicles scheduling, but also in the vessel and

other fields, and the solid technical support will be provided for the ITS research. With the wireless network augmentation, ITS application will assume a new phase of rapid development.

References

1. Feiyue, W.: Integrated intelligent control and management for urban traffic systems. In: IEEE International Conference on Intelligent Transportation System, vol. 2, pp. 1313–1317. IEEE Computer Society, Los Alamitos (2003)
2. Wei, Z., Qingling, L.(eds.): The Design and Implementation of Vehicle Monitoring and Alarm System. In: 2003 IEEE International Conference on Intelligent Transportation System, vol. 2, pp. 1158–1160 (2003)
3. Dorigo, M., Maniezzo, V., Colorni, A.: The Ant System: Optimization by a colony of cooperating agents. IEEE Transactions on Systems, Man, and Cybernetics-Part B 26(1), 29–41 (1996)
4. Dorigo, M., Gambardella, L.: Ant colonies for the traveling salesman problem. BioSystems 43(2), 73–81 (1997)
5. Li, L., Feiyue, W.: Vehicle trajectory generation for optimal driving guidance. In: 5th International Conference on Intelligent Transportation Systems, pp. 231–235 (2002)
6. Zhao, Y.: Mobile phone location determination and its impaction Intelligent Transportation Systems. lEEE Transactions on Intelligent Transportation Systems 1, 55–64 (2000)
7. Kaihua, X., Yuhua, L.: A Novel Intelligent Transportation Monitoring and Management System Based on GPRS. In: IEEE International Conference on Intelligent Transportation System, vol. 2, pp. 1654–1659 (2003)
8. Jianping, X., Jun, Z., Hebin, C.: GPS Real Time Vehicle Alarm Monitoring System Based on GPRS/CSD using the Embedded System. In: Proceedings of 6th International Conference on ITS Telecommunications 2006, vol. 1, pp. 1155–1158 (2006)

Selective Sensor Node Selection Method for Making Suitable Cluster in Filtering-Based Sensor Networks*

ByungHee Kim and Tae HoCho

School of Information and Communication Engineering, Sungkyunkwan University, Suwon 440-746, Republic of Korea
{bhkim,taecho}@ece.skku.ac.kr

Abstract. Sensor nodes, that have a limited battery power, are deployed in open and unattended environments in many sensor network applications. An adversary can compromise the deployed sensor nodes and easily inject fabricated reports into the sensor network through the compromised nodes. Filtering-based secure methods have been proposed to detect and drop the injected fabricated reports. In this scheme, the number of key partitions is important, since a correct report should have to an efficient number of different message authentication codes made by each difference key partition. We propose a selective sensor node selection method to create suitable clusters and enhance the detection power of the filtering scheme. We use a fuzzy system to enhance cluster fitness. The proposed method demonstrates sufficient resilience and energy efficiency based on the simulation results.

Keywords: Sensor network, fabricated report, clustering, filtering scheme, fuzzy system.

1 Introduction

Sensor networks are expected to interact with the physical world at an unprecedented level of universality and enable various new applications [1]. Sensor networks typically comprise a few base stations that collect the sensor readings and forward sensed information to managers and a large number of small sensor nodes that have limited processing power, small storage space, narrow bandwidth and limited energy. In many sensor network applications, sensor nodes are deployed in open and unattended environments. Sensor nodes are thus vulnerable to physical attacks potentially compromising their cryptographic keys [2]. An adversary can inject fabricated reports that are non-existent events or faked through the compromised nodes. Its goal is to spread false alarms that waste real world effort to respond. It depletes the limited energy resource of the sensor nodes [1]. Several security solutions [3-8] have been proposed to detect and drop the injected fabricated reports. One of the proposed methods is a statistically en-routing filtering scheme (SEF). SEF can filter out the injected fabricated reports before they consume significant amount of energy. The key idea of the

* This research was supported by the MKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement). (IITA-2008-C1090-0801-0028).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 636–643, 2008.
© Springer-Verlag Berlin Heidelberg 2008

proposed methods is that every sensor node verifies the validity of the event report using symmetric keys. When an interesting event occurs, a cluster head gathers a message authentication code (MAC) from the event-sensing nodes. After gathering the information, that node creates an event report using the information received. In this scheme, the number of different key partitions is important, since cluster heads do not correctly report an event when they do not receive an efficient number of MACs from cluster nodes. Therefore, every cluster region should contain the efficient number of the key different partitions.

In this paper, we propose a selective sensor node selection method for making suitable cluster regions. The proposed method can make a suitable cluster region for SEF and conserve the energy consumed. Fabricated reports can be filtered by enhancing the detection power. The effectiveness of the proposed method is shown in the simulation results.

The remainder of this paper is organized as follows. Section 2 briefly reviews a filtering scheme based on a statistically en-route filtering scheme and clustering. Section 3 details the proposed method. Section 4 reviews the simulation results. Finally, section 5 concludes the paper.

2 Background

In this section, we briefly describe SEF, clustering, and our motivation.

2.1 Statistical En-Route Filtering Scheme (SEF)

Yu, Luo, and Lu were to address the filtering-based secure method to detect and drop the injected fabricated reports. SEF assumes that the base station maintains a global key pool that is divided into multiple partitions. Each sensor node loads a small number of keys from a randomly selected partition in the global key pool before the it is deployed in a particular region. In SEF, a sensor node has certain probability of detecting fabricated reports. When an event occurs, one of the event-sensing nodes collects message authentication codes (MAC) from the other detecting nodes. That node creates an event report and forwards it to the base station. During the transmission of the event report, forwarding nodes verify the correctness of the received event report using symmetric keys.

2.2 Clustering

Many routing protocols to use sensor network are proposed to conserve sensor node energy [9-11]. Clustering is efficient in reducing energy consumed for applications that require scalability to hundreds or thousands of sensor nodes. Clustering was proposed as a useful tool to efficiently pinpoint object location [10]. Clustering can be effective when sensor nodes receive and send data to other sensor nodes.

One of the deployed sensor nodes is selected to be cluster head after the sensor node deployment. Cluster heads aggregate data of sensor node in their cluster and forward gathered data to the base station. Clustering can prolong the life time of the sensor node by reducing multiple forwarded data. Fig. 1(a) and (b) show that the multi-hop routing protocol and multi-hop routing protocol with clustering.

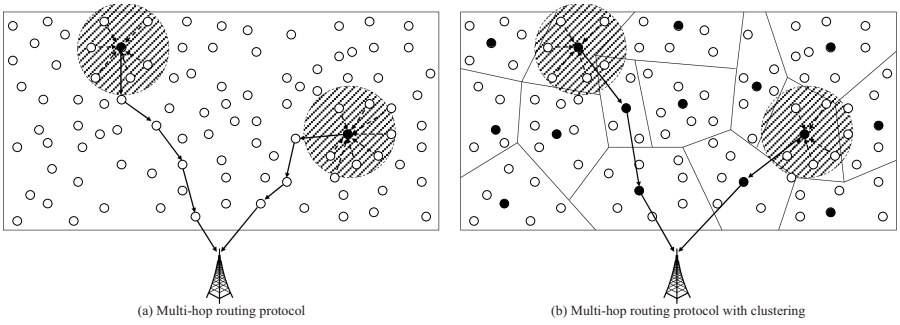


Fig. 1. Multi-hop routing protocols of in the sensor network

2.3 Motivation

Many clustering methods are proposed to conserve energy when forwarding and receiving data among sensor nodes. None of proposed methods consider filtering schemes. Sensor nodes have to select one cluster head when they detect multiple cluster heads. In SEF, if sensor nodes select a cluster head without considering the detection power, sensor node not only do not have an efficient detection power, but also do not create a correctness event.

Therefore, we should design a clustering method that guarantees efficient detection power and create suitable cluster regions.

3 Selective Selection Method Considered Fitness of the Cluster

In this section, we detail an efficient method of selective sensor node selection. We use a fuzzy system to evaluate the fitness of the sensor node with cluster heads.

3.1 Assumption

We assume that a sensor network comprises a large number of sensor nodes and is cluster-based. Sensor nodes can know their position, verify the broadcast message, and have unique identifiers. We further assume that sensor nodes are to be organized a cluster automatically on deploying in the interesting region to conserve energy and reduce transmission of duplicate data. The base station cannot be compromised by an adversary and it has a mechanism to authenticate a broadcast message.

3.2 Selective Sensor Node Selection Method Considered Fitness of Sensor Nodes

Sensor nodes are organized in clusters after deployment. When sensor nodes select a cluster head, they consider the radio weigh of a cluster head, distance, and so on. In SEF, this organization method does not guarantee sensor networks the detection power for filtering fabricated reports. We propose a selective sensor node selection method, considering fitness of sensor nodes with cluster heads, using a fuzzy system, to enhance detection power. Our proposed method has two phases: *Initial fitness phase* and *adjusted fitness phase*.

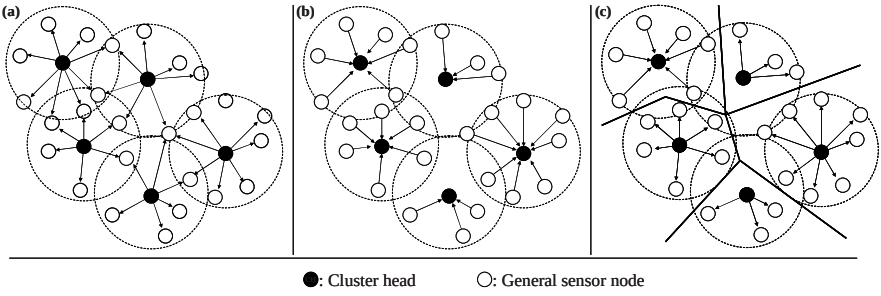


Fig. 2. The initial fitness phase to form a cluster region

Fig. 2(a), (b), and (c) show the *initial fitness phase*. In this phase, a cluster head forwards key partition information of the forwarding nodes to the near by sensor node (Fig. 2(a)). Each sensor node evaluates fitness with each cluster head and determines where it joins (Fig. 2(b)). Each sensor node joins a cluster that it has the highest fitness compared with others (Fig. 2(c)).

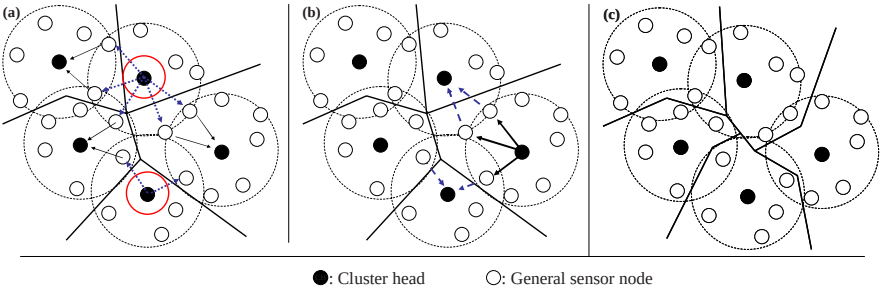


Fig. 3. The adjusted fitness phase to form a cluster region

Fig. 3 shows the *adjusted fitness phase*. This phase executes when a cluster head does not have an efficient number of sensor nodes to make an event report. In Fig. 3(a), two cluster regions in the red circle need to re-cluster, since they do not compose a correct event report. To reorganize a cluster, two cluster heads send a number of the sensor nodes in their cluster region to sensor nodes nearby (Fig. 3(b)). Sensor nodes, that receive the message, re-evaluate and revise the fitness of the cluster head (Fig. 3(c)). After revision, sensor nodes reorganize their cluster region.

3.3 FuzzyRule-Based System

We use the fuzzy system to determining the fitness of nodes for form a cluster. The fuzzy system uses three factors to determine fitness: 1) Distance from the cluster head to the base station (represented by DISTANCE), 2) Weigh of partition information (represented by WPI), and 3) The number of sensor nodes in a cluster region (represented by NSC). In our proposed method, fuzzy sets are used to determine fitness of sensor nodes at the *initial fitness phase* and *adjusted fitness phase*.

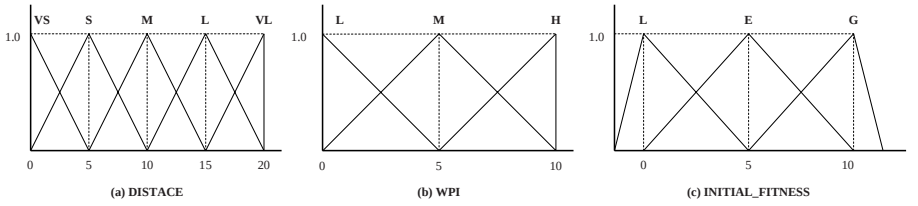


Fig. 4. Fuzzy membership functions to determining initial fitness

Fig. 4(a), (b), and (c) illustrate two input and one output membership functions (represented by INITIAL_FITNESS) in the *initial fitness phase*.

In the *adjusted fitness phase*, the result of the *initial fitness phase* (represented by PFV) is used to determine fitness with cluster heads. Fig. 5(a), (b), and (c) illustrate the membership functions of the two fuzzy logic input parameters used in the *adjusted fitness phase* and one output membership functions (represented by ADJUSTED_FITNESS).

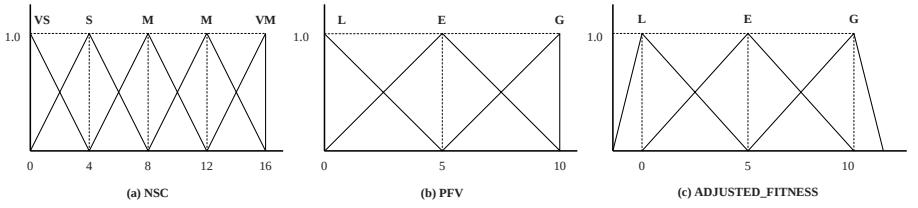


Fig. 5. Fuzzy membership functions to determining adjusted fitness

We defined fuzzy if-then rules to use the fuzzy membership function in the fuzzy system. Table 1 shows some rules of the fuzzy rule-based system.

Table 1. Fuzzy if-then rules

Rule No.	INITIAL FITNESS PHASE			ADJUSTED FITNESS PHASE		
	IF		THEN	IF		THEN
	DISTANCE	WIP	INITIAL_FITNESS	NSC	PFV	ADJUSTED_FITNESS
1	VS	L	L	VF	VG	VL
2	S	M	L	F	G	VL
3	M	H	G	M	E	E
4	L	L	L	M	L	G
5	VL	M	E	VM	VL	G

4 Simulation Evaluation

We compare the clustering method, distance-based cluster determining method (represented by DCDM) and fuzzy-based cluster determining method (represented by FCDM). In DCDM, sensor nodes select a cluster head by considering the distance from sensor node to cluster head. FCDM uses the fuzzy system to organize a cluster

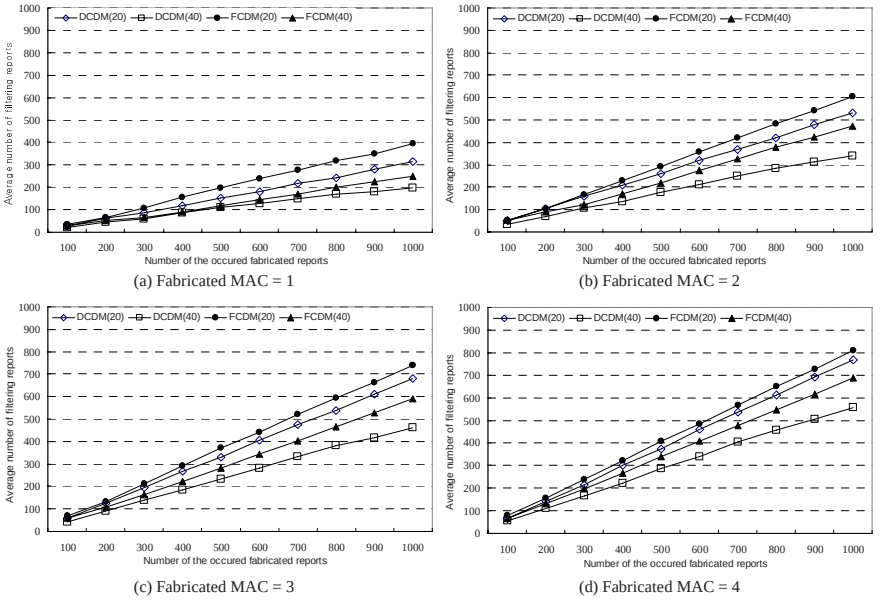
from detecting cluster heads. In FCDM, sensor nodes use fitness calculated by the fuzzy system to select a cluster head.

4.1 Simulation Environment

To show the effectiveness of the proposed method, we compare the proposed method with DCDM about filtering ratio and travel hops of fabricated reports in SEF. We use a field size of 500×500 m^2, where cluster heads are uniformly distributed and 2600 general sensor nodes are randomly distributed. The base station is located at the end of the field. The global key pool is divided into 20 and 40 key partitions to compare the detection power following the number of key partitions. Each key partition has 100 secure keys and each sensor node randomly selects 50 keys from one of key partitions.

4.2 Simulation Results

We compare the performance of our scheme with that of the other method by simulation. The simulation results, that are the number of filtered reports and travel hops of fabricated reports, show that our proposed method is more efficient than the comparable methods. We also consider the number of key partitions in the global key pool and simulated methods when the global key pool is divided into 20 and 40 key partitions. Based on the number of key partitions, FCDC is divided FCDC (20) (filled circle) and FCDC (40) (filled triangle), and DCDC is divided DCDC (20) (empty diamond) and DCDC (40) (empty rectangle).



(a) Fabricated MAC = 1 (b) Fabricated MAC = 2

(c) Fabricated MAC = 3 (d) Fabricated MAC = 4

Fig. 6. The filtering ratio of fabricated reports when the number of key partitions in the global key pool is 20 and 40

Fig. 6(a), (b), (c), and (d) show the number of filtered reports when fabricated MACs are one, two, three, and four. Our proposed method is more efficient than DCDC when the number of partitions is the same.

Fig. 7(a), (b), (c), and (d) show the average travel hops of fabricated reports. As shown in the figure, the proposed method is more efficient than other methods. When the number of the fabricated MACs is one, FCDC(20) can reduce travel hops to 86.71% compared whit DCDC(20) and FCDC(40) reduce it to 92.92%. FCDC(20) can reduce travel hops to 75.59% compared with DCDC(20) and FCDC(40) reduce it to 74.99% compared with DCDC(40) when the fabricated MACs is four.

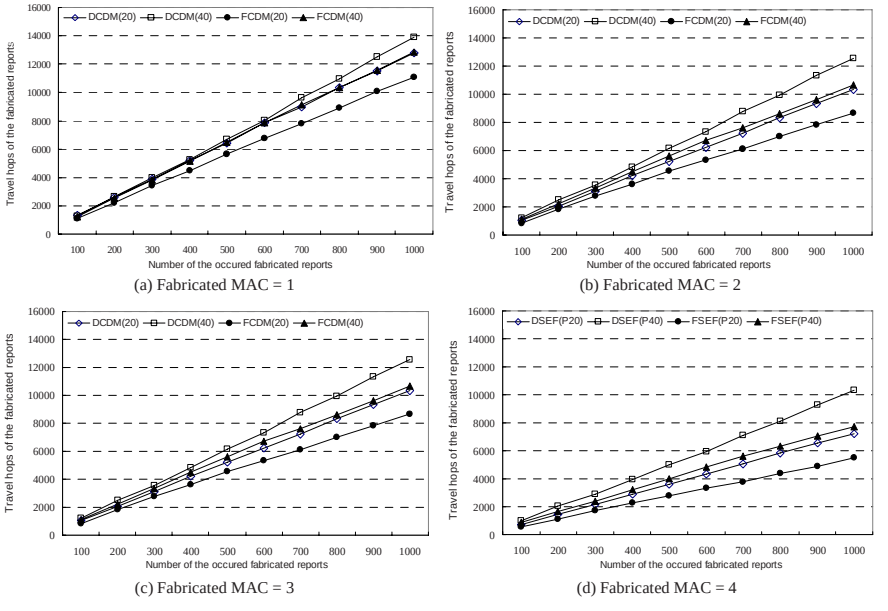


Fig. 7. The travel hops of fabricated reports when the number of key partitions in the global key pool is 20 and 40

As demonstrated by the simulation results, the detection power of the proposed method is more efficient than for DCDC. We further know that a large number of key partitions have more efficient detection power than a few key partitions when many fabricated MACs are in an event report. When the fabricated MACs are few, a few smaller partitions is more efficient than a larger key partition.

5 Conclusion and Future Work

In this paper, we have proposed the selective sensor node s method to create clusters to enhance detection power in sensor networks. A fuzzy rule-base system is exploited to evaluate the fitness of clustering adaptation. The fuzzy system uses the partition information of forwarding nodes, distance from a cluster to the base station, and

number of sensor nodes in a cluster to determine fitness value. The proposed method can enhance detection power. The effectiveness of the proposed method is demonstrated by the simulation results.

The proposed method is applied to SEF for simulation. Future research will apply other filtering schemes and run additional simulation with input factors not covered in this work.

References

1. Al-Karaki, J.N., Kamal, A.E.: Routing techniques in wireless sensor networks: a survey. IEE Wirel. Commun. 11(6), 6–28 (2004)
2. Li, F., Wu, J.: A Probabilistic Voting-Based Filtering Scheme in Wireless Sensor Networks. In: The International Wireless Communications and Mobile Computing Conference, pp. 27–32 (2006)
3. Yang, H., Lu, S.: Commutative Cipher Based En-Route Filtering in Wireless Sensor Networks. In: The IEEE Vehicular Technology Conference, pp. 1223–1227 (2003)
4. Ye, F., Luo, H., Lu, S.: Statistical En-Route Filtering of Injected False Data in Sensor Networks. IEEE J. Sel. Area Comm. 23(4), 839–850 (2005)
5. Zhu, S., Setia, S., Jajodia, S., Ning, P.: An Interleaved Hop-by-Hop Authentication Scheme for Filtering of Injected False Data in Sensor Networks. In: The IEEE Symposium on Security and Privacy, pp. 259–271 (2004)
6. Lee, H.Y., Cho, T.H.: Key Inheritance-Based False Data Filtering Scheme in Wireless Sensor Networks. In: Madria, S.K., Claypool, K.T., Kannan, R., Uppuluri, P., Gore, M.M. (eds.) ICDCIT 2006. LNCS, vol. 4317, pp. 116–127. Springer, Heidelberg (2006)
7. Lee, H.Y., Cho, T.H.: Fuzzy Adaptive Selection of Filtering Schemes for Energy Saving in Sensor Networks. IEICE Trans. Commun. E90–B(12), 3346–3353 (2007)
8. Yu, Z., Guan, Y.: A Dynamic En-route Scheme for Filtering False Data Injection in Wireless Sensor Networks. In: Proc. Of SenSys., pp. 294–295 (2005)
9. Zhang, W., Cao, G.: Group Rekeying for Filtering False Data in Sensor Networks: A Predistribution and Local Collaboration-Based Approach. In: INFOCOM 2005, pp. 503–514 (2005)
10. Huang, C.C., Guo, M.H.: Weight-Based Clustering Multicast Routing Protocol for Mobile Ad Hoc Networks. Internet Protocol Tec. 1(1), 10–18 (2003)
11. Ossama, Y., Sonia, F.: HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks. IEEE Trans. Mobile Com. 3(4), 366–379 (2004)

Formal Description and Verification of Web Service Composition Based on OOPN

Jindian Su[1], Shanshan Yu[2], and Heqing Guo[1]

[1] College of Computer Science and Enginerring, South China University of Technology, 510006 Guangzhou, China
[2] School of Information and Technology, Computer Sciecent, Sun Yat-sen Unversity, 510275 Guangzhou, China
SuJD@scut.edu.cn, Yushsh@mail.sysu.edu.cn, Guozou@scut.edu.cn

Abstract. To overcome the shortcomings of unable to demonstrate the characteristics of encapsulation, reusability and message-driven when using Petri nets to model Web service composition, this paper proposes object oriented Petri nets (OOPN) as the formal description and verification tool for Web service composition. Some basic concepts of OOPN, including class net, inheritance, object net and object net system, are given in the paper respectively and the mechanisms of "Gate" and "Port" are used to describe message transfer and synchronization relationships among objects. The OOPN expressions of various basic control flow structure and message synchronization relationships of Web service composition are also stated. Finally, we give a case study to show that our proposed method can more naturally and accurately demonstrate the encapsulation and message-driven characteristics of Web services.

Keywords: Formal Description, Web Service Composition, Object Oriented Petri Nets.

1 Introduction

With the increasing complexity of business requirements and logical rules, plus the louse-coupled, distributed and flexible characteristics of Web services, the possibility of errors in Web service composition (WSC) is greatly increased, which makes traditional manual analysis methods become unsuitable. As a result, many scholars tried to propose formal methods, such as Pi calculus, Petri nets, or Finite State Machine, to build the formal description and verification models of WSC. Among these formal methods, Petri nets, as a graphical and mathematical modeling tool, is quite suitable Web services[1～4]. It can provide accurate semantic explanations and qualitative or quantitative analyses about various dependent relationships among services. But Petri nets can't directly and naturally express some important characteristics of Web services, such as encapsulation, reusability and message-driven.

Object oriented Petri nets (OOPN), as a kind of high level Petri nets, takes full advantages of object oriented (OO) theory and the advantages of Petri nets. It incorporates various techniques of OO, such as encapsulation, inheritance and message driven, into coloured Petri nets (CPN) and becomes a very suitable for WSC.

The remainders of this paper are organized as follows: Section 2 discusses some related work. The formal definitions of OOPN are presented in section 3. In section 4,

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 644–652, 2008.
© Springer-Verlag Berlin Heidelberg 2008

we explain how to use OOPN to build WSC models and also discuss the reachability and boundness properties of OOPN. In section 5, we use a case study to show the applications of OOPN. The conclusions are given in section 6.

2 Related Work

Over the past decades, the combinations of Petri nets with OO theory have been extensively studied and some OOPN definitions, such as PROTOB[5], OPNets[6], POT[7] and OOCPN[8], were proposed consequently. The main ideas of PROTOB and OPNets are to divide a system into many objects, each of which is depicted by a subnet and the system's behaviors are described with the help of message transfer mechanism. OOCPN is an extension of hierarchical coloured Petri nets (HCPN), and the tokens that flow in the nets are objects. Both OOCPN and POT are object-based instead of object-oriented and don't base on inheritance and encapsulation, which as a result can't decrease the complexity of formal models. COOPN/2 used extended Petri nets to describe objects and expressed both abstract and concrete aspects of systems by putting emphasis on the descriptions of concurrent and abstract data types.

[9] proposed an architecture description language based on OOPN and used it to model multi agent system. [10] proposed an OO modular Petri nets (OOMPNet) for modeling service oriented applications. In OOMPNet, different object nets were fused through shared nodes and object nets could be tokens that flowed in another net. Both [9] and [10] neglected inheritance mechanism and can't express message-driven characteristics of OO. [11] used OOPN to discuss the formal semantics of WSC operations and control flow models. But the usage of port places in control flow models deserves considerations and the authors didn't provided any discussions about inheritance or message synchronization relationships. [12] proposed a technique to transform an OO design into HCPN models with abstract node approach, but they didn't discuss how to model the structure and behaviors of systems with OOPN.

Compared with COOPN/2 and POT, the tokens that flow in our proposed OOPN are messages instead of objects, and the internal structure and external structure of objects are separated by message interfaces.

3 Object Oriented Petri Nets

OOPN provides a sound mathematical foundation for describing objects and their message transfer and synchronization relationships. An object, composed of external structure and internal structure, is depicted by an object net and message interfaces are the only way of interaction. Different object nets might work together to finish business tasks and as a result form into an object net system.

3.1 Class Net－*Cnet*

Class is an encapsulation body of properties and operations. Its internal properties and states can be mapped to places and operations can be mapped to behavior transitions. A class can be expressed as an OOPN subnet, which is enclosed with a square box to

express the meanings of encapsulation. Each class net has two parts, external structure and internal structure. The formal definitions of class net are shown as follows:

Definition 1 (Class Net). A class net is a 9-tuple $Cnet=\{\Sigma,P,T,IM,OM,K,C,G,F\}$, where

1) Σ is a finite set of color types of all places, known as color set.
2) $P=Pa\bigcup Ps$ is a set of places, where Pa and Ps are the place sets of static properties and states respectively.
3) T is a finite set of behavior transitions.
4) IM and OM are two place sets of input and output messages respectively, also called message interfaces. They are the only places for the class to send or receive messages and are both connected with input or output arcs. $IM=\{im_i,i=1,2,\cdots,n\}$ and $OM=\{om_j,j=1,2,\cdots,n\}$ should appear in pairs. The values of input or output messages are hide in the tokens of IM and OM. If several different objects send messages to an object at the same time, then these messages will automatically form into a queue in IM of the target object. There is $\neg\exists t\in T:<t,im_i>\in F$, $\neg\exists t\in T:<om_j,t>\in F$.
5) $K:P\bigcup IM\bigcup OM\rightarrow N$ is a capacity function defined on places and has $\forall p\in IM\bigcup OM:K(p)=k$, where k ($k\geq 1$) is an integer.
6) $C:P\rightarrow\Sigma$ is a color function that denotes the types of each place in *Cnet*.
7) $G:T\rightarrow G(t)$ is a guard function defined on T and has $\forall t\in T$: $[Type(G(t))=bool\wedge Type(Var(G(t)))\subseteq\Sigma]$.
8) $F=G\bigcup H$ is a set of directed arcs. $G\subseteq(P\times T)\bigcup(T\times P)$ is the set of flow relationships between transitions and places, and $H\subseteq(IM\times T)\bigcup(T\times OM)$ is the set of flow relationships between message interfaces and operations.

The graphical expression of *Cnet* can be seen in Fig.1.

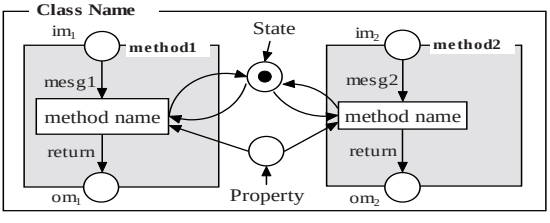


Fig. 1. Graphical Expression of *Cnet*

Inheritance is a kind of partial order relationships defined on classes and defined as:

Definition 2 (Inheritance). If $Cnet_1=\{\Sigma_1,P_1,T_1,IM_1,OM_1,K_1,C_1,G_1,F_1\}$ and $Cnet_2=\{\Sigma_2,P_2,T_2,IM_2,OM_2,K_2,C_2,G_2,F_2\}$ are two different class nets and $Cnet_2$ is a subnet of $Cnet_1$, denoted as $Cnet_2\prec Cnet_1$, when they satisfy:

1) $P_2 \subseteq P_1$; 2) $T_2 \subseteq T_1$;3) $IM_2 \subseteq IM_1$, $OM_2 \subseteq OM_1$;

4) $IM_2 \times T_2 \subseteq IM_1 \times T_1$, $T_2 \times OM_2 \subseteq T_1 \times OM_1$

3.2 Object Net and Object Net System

An object net (*Onet*) can be viewed as a *Cnet* with initial markings and is defined as:

Definition 3 (*Onet*). An object net is a 2-tuple $Onet = \{Cnet, M_0\}$, where

1) *Cnet* is a class net; 2) M_0 is the initial markings of *Cnet*

The graphical expression of *Onet* is shown in Fig.2.

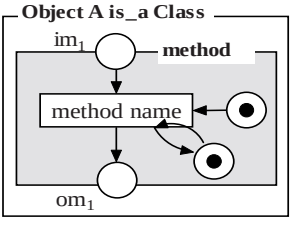


Fig. 2. Graphical Expression of *Onet*

Each *Onet* is an independent entity and has its internal state places and behavior transitions. But *Onet* can also cooperate with others through message transfer in order to finish tasks, which forms into an object net system, called *Onets*. In *Onets*, objects communicate with each other through message interfaces connected by gate transitions and their message synchronization relationships can be depicted by port places. Gate transitions and port places constitute the control structure of *Onets*.

Definition 4 (*Gate*). *Gate* is a kind of special transitions used to describe message transfer relationships among objects. It connects an *OM* places with an *IM* places.

$$\forall t \in Gate : {}^{\bullet}t \in OM \wedge t^{\bullet} \in IM$$

Definition 5 (*Port*). *Port* is a kind of special places used to express message synchronization relationships among objects and exists between two gate transitions.

$$\forall p \in Port : {}^{\bullet}p \in Gate \wedge p^{\bullet} \in Gate$$

As an object provides services through its operations, so gate transitions and port places can also be applied to the different operations of the same object.

Definition 6 (Object Net System). An object net system is a 9-tuple $Onets = \{\Sigma, O, Gate, Port, G, IM, OM, F, M_0\}$, where

1) Σ is a color set of place types in *Onets*.
2) $O = \{o_i, i = 1,2,...,N\}$ is the set of all objects in *Onets* and each object o_i can be an *Onet* or an *Onets*.
3) $Gate = \{gate_{ij}, i, j = 1,2,...,N\}$ is a finite set of gates, where
   $Gate \subseteq (OM \times Gate) \bigcup (Gate \times IM)$, $im_i \times gate_{ij} \in F$, $gate_{ij} \times om_j \in F$

4) $Port = \{port_{ijmn}, i, j, m, n = 1,2,...,N\}$ is a set of port places and their capacity values are 1 by default.
5) $G: Gate \rightarrow G(t)$ is a guard function defined on gates, where $\forall t \in Gate$: $[Type(G(t)) = bool \wedge Type(Var(G(t))) \subseteq \Sigma]$.
6) *IM* and *OM* are the sets of input and output message interfaces of all objects in *Onets* respectively.
7) $F \subseteq R \bigcup S$ is a finite set of flow relationships. $R = \{Rij, i, j = 1,2,...,N, i \neq j\}$ is the flow relationship set between message sender O_i and message receiver O_j. For each $R_{ij} = \{om_i, gate_{ij}, im_j\}$, it has $om_i \in OM$ and $im_j \in IM$. $S = \{s_{ijmn}, i, j, m, n = 1,2,...,N, i \neq j \neq m \neq n\}$ is a set of message synchronization relationships and $s_{ijmn} = \{gate_{ij}, port_{ijmn}, gate_{mn}\}$.
8) M_0 is the initial markings of *Onets*.

4 OOPN-Based WSC Model

Web services can be viewed as autonomous objects deployed in the Web and might cooperate with others through message transfer. As a result, each atomic service can be mapped to an *Onet*, and each WSC can be described by an *Onets*, which is composed of different service objects and the corresponding dependent relationships.

4.1 OOPN Descriptions of Control Structure in WSC

The common control structure in WSC usually includes sequential, selective, concurrent, loop and message synchronization.

Suppose S_1, S_2 and S_3 are three different atomic Web services, each of which only includes one operation, denoted as op_1, op_2 and op_3 respectively. The algebra operator descriptions of these control structure can be seen as: $S = X|S_1 \bullet S_2|S_1 \oplus S_2|S_1 \otimes S_2|nS_1|S_1||_{s_3} S_2$

1) X stands for an empty service (a service without executing any operation).
2) $S_1 \bullet S_2$ means that S_1 and S_2 should execute sequentially and $\bullet$ is the sequential operator. (see Fig.3)

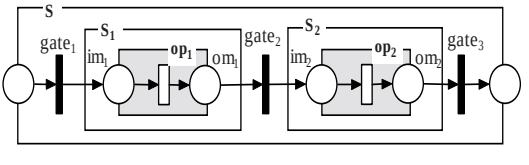


Fig. 3. OOPN expression of $S_1 \bullet S_2$

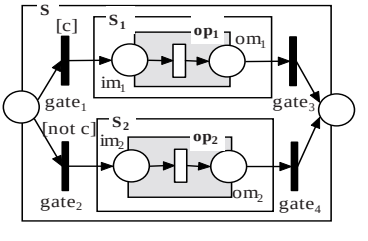


Fig. 4. OOPN expression $S_1 \oplus S_2$

3) $S_1 \oplus S_2$ means that only S_1 or S_2 will be selected for execution and they can't execute concurrently. $\oplus$ is the selective operator .(see Fig.4)

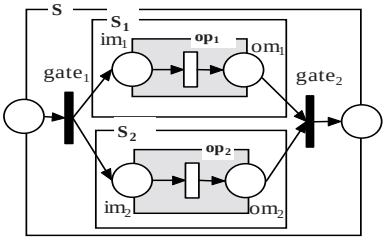


Fig. 5. OOPN expression $S_1 \otimes S_2$

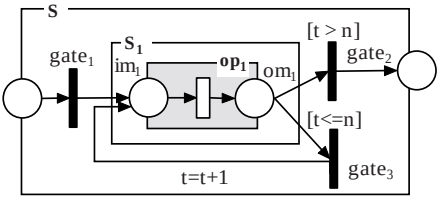


Fig. 6. OOPN expression of nS_1

4) $S_1 \otimes S_2$ means that S_1 and S_2 can execute concurrently and $\otimes$ is the concurrent operator. (see Fig.5)
5) nS_1 means that to execute S_1 for n times. (see Fig.6)
6) $S_1 ||_{s3} S_2$ means that message synchronization relationships occur when S_1 and S_2 send message requests to S_3 at the same time. $||_{s3}$ is the synchronization operator and $port_1$ is a port place. (see Fig.7)

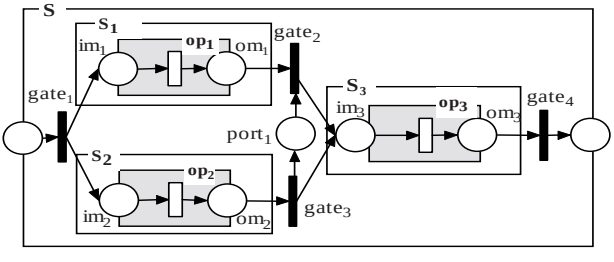


Fig. 7. OOPN expression $S_1 ||_{s3} S_2$

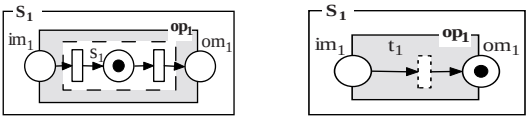
4.2 Analysis of Reachability and Boundness

The basic properties of Petri nets usually include reachability, boundness, liveness, integration and advancement, et al. Due to the limited spaces, we only discuss reachability and boundness of OOPN.

(1) Reachability. In *Onets*, message interfaces separate object's internal structure from its external structure. As a result, they tend to cause the embarrassment situations of unable to determine the states of *Onets* when analyzing *Onets*.

For example, in Fig.8 (a), S_1 is processing a request message. Then, if we use the internal states of S_1 to express the current states of *Onets*, it will destroy the encapsulation of S_1 according to OO theory. However, if we don't do that, it will mean that

Onets can be stateless at sometime. To solve this problem, we can use an immediate transition t_1 in Fig.8 (b) to denote the internal structure of S_1 according to the encapsulation of OO theory. It means that we only concern about the input and output states of objects instead of their internal states when analyzing *Onets*,



(a). S_1's internal states (b). Abstraction of S_1's internal States

Fig. 8. Abstraction of S_1's internal states

(2) Boundness. In theory, other objects can send messages to the *IM* places of another object unlimitedly. As a result, *Onet* and *Onets* become unbounded. However, in practical applications, message queue buffer resources are usually limited, so we can put capacity restrictions on *IM* and *OM* and make OOPN become a bounded net.

5 A Case Study

Suppose a user needs to develop an itinerary before attending a meeting from city A to city B, which includes transport types and hotel booking services. There are some restrictions on choosing vehicles. If the traveling time by train from A to B is less than or equal to six hours, he will travel by train; Otherwise, he will choose airplane.

Then, the algebra operator expression of WSC is:

$$Onets = WS_1(op_1) \bullet (WS_2(op_2 \bullet op_3) \oplus WS_3(op_4 \bullet op_5)) \bullet WS_4(op_6) \bullet WS_4(op_7)$$

The OOPN expression of *Onets* and its reachability graph are given in Fig.9 and Fig.10. From Fig.9, we can see that the OOPN-based WSC model can accurately and

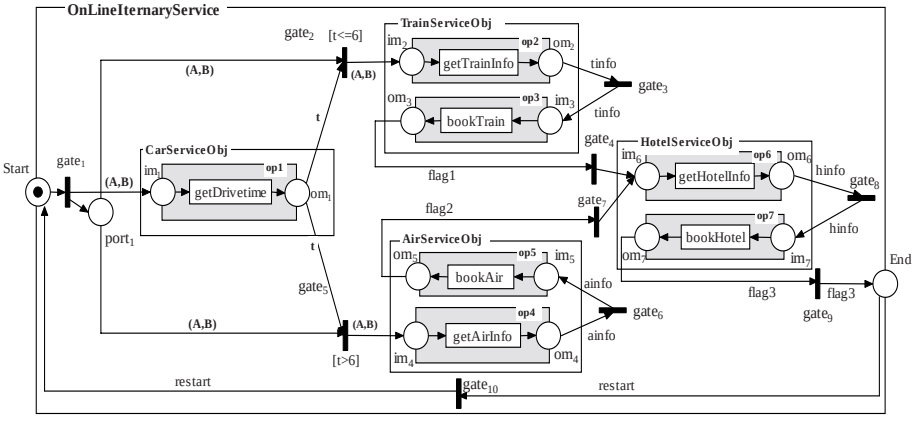


Fig. 9. OOPN model of online itinerary example

naturally demonstrate the encapsulation and message-driven characteristics of Web services. Compared with the formal models in [1, 3], our OOPN method is more suitable for Web services.

By using reachability graph RG(*Onets*) in Fig.10 to analyze the above example, we can learn that *Onets* has the following properties:

(1) Reachability. Obviously, for any M_i, $M_j \in R(M_0)$, ($i<j$), M_j is reachable from M_i. $R(M_0)$ is the reachable marking set of *Onets*.

(2) Boundness. *Onets* is bounded and the upper limit number of tokens depends on the capacities of *IM* and *OM*.

(3) Liveness. *Onets* is live, because any directed path from M_0 can reach a strong graph eventually, in which each transition $t \in T \cup Gate$ at least belongs to the label of some directed arc and can be fired by some transition list.

(4) Integration. All states in *Onets* are reachable from M_0 and will eventually get to the final state M_{15}.

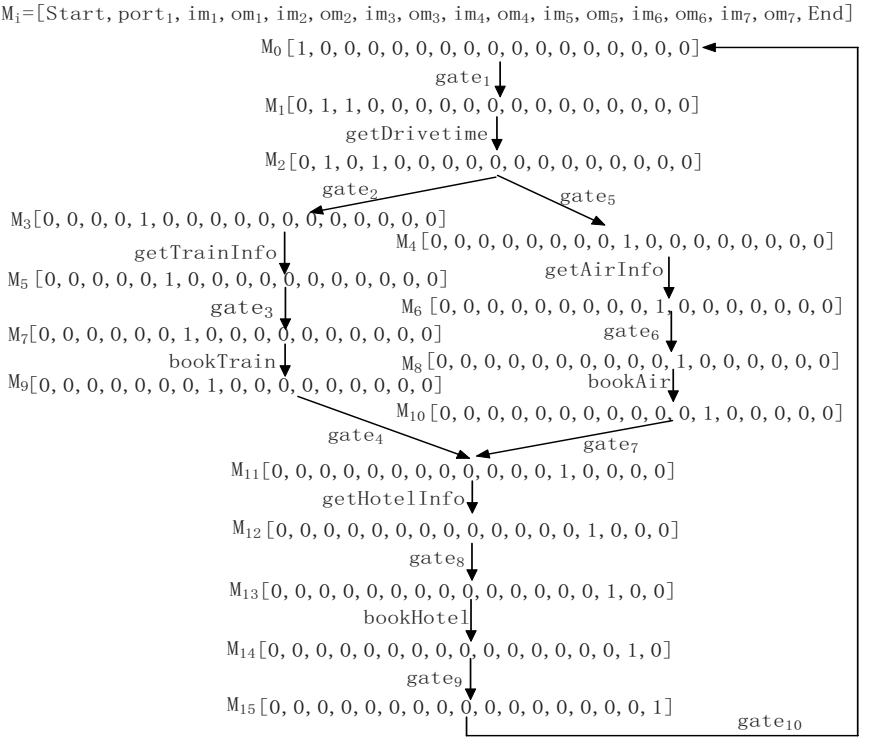


Fig. 10. Reachability Graph RG(*Onets*) of Fig. 9

6 Conclusions

Formal description and verification of WSC is a research hotspot in the field of Web services. By keeping the original semantics and properties of ordinary Petri nets, the

OOPN proposed in this paper takes full advantages of OO theory and uses Onets to describe different objects and their various kinds of dependent relationships caused by message transfer. The basic control structure and message synchronization relationships are also discussed in this paper. Finally, we use a case study to illuminate that the OOPN-based Web service composition model can well demonstrate the characteristics of encapsulation and message-driven of Web services and has a sound mathematical foundation.

Nevertheless, we just give some basic definitions of OOPN and still lack of in-depth analysis. In the future, we intend to provide more detail discussion about the structural properties and dynamic behavior characteristics of OOPN.

Acknowledgments. This research was supported by Guangdong Software and Application Technology Laboratory (GDSATL).

References

1. Guo, Y.B., Du, Y.Y., Xi, J.Q.: A CP-Net Model and Operation Properties for Web Service Composition. Chinese Journal of Computer 29, 1067–1075 (2006)
2. Tang, X.F., Jiang, C.J., Wang, Z.J.: A Petri Net-based Semantic Web Service Automatic Composition Method. Journal of Software 19, 2991–3000 (2007)
3. Zhang, P.Y., Huang, B., Sun, Y.M.: Petri-Net-Based Description and Verification of Web Services Composition Model. Journal of System Simulation 19, 2872–2876 (2007)
4. Li, J.X.: Research on Model for Web Service Composition Based on Extended Colored Petri Net. Ph.d Thesis, Graduate University of Chinese Academy of Science, Beijing (2006)
5. Baldassari, M., Bruno, G.: PROTOB: An Object Oriented Methodology for Developing Discrete Event Dynamic Systems. Computer Languages 16, 39–63 (1991)
6. Kyu, L.Y.: OPNets:An Object-Oriented High Level Petri Net Model for Real-Time System Modeling. Journal of Systems and Software 20, 69–86 (1993)
7. Engelfriet, J., Leih, G., Rozenberg, G.: Net Based Description of Parallel Object-Based Systems, or POTs and POPs. In: Proc. Workshop Foundations of Object-Oriented Languages (FOOL 1990), pp. 229–273. Springer, Heidelberg (1991)
8. Buchs, D., Guelfi, N.: A Formal Specification Framework for Object-Oriented Distributed Systems. IEEE Transactions on Software Engineering 26, 432–454 (2000)
9. Yu, Z.H., Li, Z.W.: Architecture Description Language Based on Object-Oriented Petri Nets for Multi-agent Systems. In: Proceedings of 2005 IEEE International Conference on Networking, Sensing and Control, pp. 256–260. IEEE Computer Society, Los Alamitos (2005)
10. Wang, C.H., Wang, F.J.: An Object-Oriented Modular Petri Nets for Modeling Service Oriented Applications. In: Proceedings of 31st Annual International Computer Software and Applications Conference, pp. 479–486. IEEE Computer Society, Los Alamitos (2007)
11. Tao, X.F., Sun, J.: Web Service Composition based on Object-Oriented Petri Net. Computer Applications 25, 1424–1426 (2005)
12. Bauskar, B.E., Mikolajczak, B.: Abstract Node Method for Integration of Object Oriented Design with Colored Petri nets. In: Proceedings of the Third International Conference on Information Technology: New Generations (ITNG 2006), pp. 680–687. IEEE Computer Society, Los Alamitos (2006)

Fuzzy Logic Control-Based Load Balancing Agent for Distributed RFID Systems

Sung Ho Jang and Jong Sik Lee

School of Information Engineering, Inha University,
Incheon 402-751, Korea
ho7809@hanmail.net, jslee@inha.ac.kr

Abstract. The current RFID system is based on a distributed computing system to process a great volume of tag data. Therefore, RFID M/W should provide load balancing for consistent and high system performance. Load balancing is a technique to solve workload imbalance among edge M/Ws and improve scalability of RFID systems. In this paper, we propose a fuzzy logic control-based load balancing agent for distributed RFID systems. The agent predicts data volume of an edge M/W and evaluates workload of the edge M/W by fuzzy logic control. And, the agent acquires workload information of an edge M/W group to which the edge M/W belongs. Also, the agent adjusts over and under-loaded workload according to a dynamic RFID load balancing algorithm. Experimental results demonstrates that the fuzzy logic control-based load balancing agent provides 95% and 6.5% reductions in overload time and processor usage and 13.7% and 12.2% improvements in throughput and turn around time as compared the mobile agent-based load balancing system.

Keywords: Fuzzy Logic, Intelligent Agent, Load Balancing, RFID Middleware.

1 Introduction

RFID technology [1] is a solution to realize ubiquitous and intelligent computing. Recently, the application of RFID technology has been extended by participation and concern of large companies like Wal-Mart and Tesco. The volume of RFID data, which should be processed on RFID M/W in real-time, has rapidly increased by the diffusion of RFID applications. A large number of tag data observed by readers is difficult to be solved by a host computer equipped with RFID M/W. The data volume of a small-scale RFID application hardly exceeds computing power of a host computer, while the data volume of a large-scale RFID application employed by major enterprises does not. For example, in case of Wal-Mart, about 7 terabyte data will be created on its RFID trial system every day [2]. Therefore, in order to process a large amount of RFID data, most RFID developers and researchers have conformed to the EPCglobal Architecture Framework Specification established by EPCglobal [3]. In the specification, a RFID system is based on the distributed system structure, several edge M/Ws included. But, these edge M/Ws have quite different and restricted computing performance. And, the volume of RFID data processed on each edge M/W is flexible. If the volume of RFID data allocated to a specific edge M/W is too much to

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 653–660, 2008.
© Springer-Verlag Berlin Heidelberg 2008

process, the edge M/W will probably generate data loss and response delay. Therefore, load balancing is required to solve workload imbalance among edge M/Ws.

In this paper, we propose the fuzzy logic control-based intelligent agent for effective RFID load balancing. The proposed agent provides the ability to gather workload information of an edge M/W group and reallocate system workload. Also, the proposed agent applies a fuzzy logic controller to evaluate workload of an edge M/W and executes dynamic load balancing, which can enhance system scalability and flexibility with high throughput.

The rest of this paper is as follows. Section 2 presents existing RFID load balancing methods and discusses their weakness. Section 3 describes the proposed fuzzy logic control-based load balancing agent. Section 4 presents the experimental results. Finally, the conclusion is given in Section 5.

2 Related Works

In a distributed RFID system, system performance of an edge M/W is degraded when tag data observed from readers is overloaded in the edge M/W. And, it leads the edge M/W to run into a breakdown. Therefore, RFID load balancing [4] is necessary to provide consistent system performance and prevent concentration of workload. Some researches were proposed for effective RFID load balancing.

Park et al proposed the RFID load balancing method using connection pool in [5]. In this work, a RFID system is required to provide a mass storage device called connection pool, which stores and manages a tremendous amount of tag data. Because, tag data from readers should be transferred to edge M/Ws through connection pool, takes charge of tag data distribution. Dong et al proposed the localized load balancing method for large-scale RFID systems in [6]. In the method, each reader computes energy cost of its edges and load balancing is executed by the energy cost. These existing load balancing methods are unreasonable to be applied to the current RFID system. They have to be adapted into a specific distributed system. In [7], Cui and Chae proposed the mobile agent-based load balancing system for distributed RFID systems. LGA (Load Gathering Agent) and RLBA (RFID Load Balancing Agent) were designed to gather global workload of RFID M/Ws and execute RFID load balancing. This system relocates readers to edge M/Ws frequently. However, reader relocation causes performance degradation of an edge M/W. The network load of overloaded edge M/Ws is deadly increased every reader relocation, due to socket connections between readers with the edge M/Ws. Also, Cui and Chae did not explain how to evaluate workload of an edge M/W.

Therefore, this paper describes the process of workload evaluation and proposes a load balancing approach specialized for a distributed RFID system. The proposed approach is based on the intelligent agent system, which can solve technical problems of distributed middleware management. For RFID load balancing, we don't relocate readers of an edge M/W to other edge M/Ws, but reallocates workload of the edge M/W to other edge M/Ws. This paper also applies the proposed load balancing approach to a distributed RFID system on the simulation environment [8].

3 Fuzzy Logic Control-Based Load Balancing Agent

The architecture of a general RFID system is composed of an integration M/W, edge M/Ws, client applications, and physical readers. An integration M/W manages several edge M/W groups and processes data received from them. And, the integration M/W communicates with client applications, such as ERP (Enterprise Resource Planning) and WMS (Warehouse Management System), and provides information inquired from the client applications. Also, the integration M/W includes IS (Information Server) to share RFID tag information and business information related with client applications. Both IS and applications define services using web service and exchange service results in the XML format through a designated URI. An edge M/W group comprises some edge M/Ws located within a local area, $EMG = \{em_1, em_2, em_3, \ldots, em_n\}$. Each edge M/W contains a set of readers, $R(em) = \{r_1, r_2, r_3, \ldots, r_m\}$. All readers are connected to only one edge M/W and belong to a set of all readers connected to an edge M/W group, $R(EMG)$. RFID tags attached to different objects are observed by readers. Readers observe data of tags placed in their vicinity and send it to an edge M/W. An edge M/W manages not only tag data but also readers. However, we assume that reader registration and deletion are not executed except in the case of initializing an edge M/W. Because, physical relocation of readers does not occur in the practical working environment and dissipates communication times and computing resources for socket connection. Therefore, the gravest task of an edge M/W is data aggregation and filtering. And, work load of an edge M/W can be defined as follows.

$$WL(em) = \sum_{i=1}^{m} CacheData(r_i) + \sum_{i=1}^{m} TagData(r_i), \forall r_i \in R(em) \tag{1}$$

In the equation (1), $WL(em)$ is work load of an edge M/W and $R(em)$ is a set of readers connected to the edge M/W. $CacheData(r)$ indicates temporary data of a reader for tag data filtering. $TagData(r)$ indicates data received from the reader in real time. $CacheData(r)$ and $TagData(r)$ are obtained from the equation (2) and (3), where $TempRead(t)$ denotes the previous tag readings stored in cache memory and $TagRead(t)$ denotes the current observed tag readings and $T(r)$ is a set of all tags observed by the reader.

$$CacheData(r) = \sum_{k=1}^{m} TempRead(t_k), \forall t_k \in T(r) \tag{2}$$

$$TagData(r) = \sum_{k=1}^{n} TagRead(t_k), \forall t_k \in T(r) \tag{3}$$

The fuzzy logic control-based load balancing agent, located in an edge M/W, controls data stream from readers and estimates workload assigned to the edge M/W. If the workload is not suitable for system performance of an edge M/W, the agent adjusts the volume of tag and cache data according to the dynamic RFID load balancing algorithm.

3.1 Estimation of Work Load

As mentioned in the previous section, workload of an edge M/W can be expressed as the sum of tag and cache data. But, if we apply the current workload measured from an edge M/W to RFID load balancing, the following problems can be encountered. Firstly, when the current workload of an edge M/W exceeds computing power of the edge M/W, data loss or delay is more likely to be generated during adjusting workload. Secondly, a host computer equipped with an edge M/W can be broken down by over loaded workload. Lastly, load balancing occurs too often due to the wide range of fluctuation in measured cache and tag data. To solve these problems, we predict the volume of cache and tag data by the equation (4) which is based on Brown's double smoothing method [9]. In this equation, P is the predicted volume of data, M is the current volume of data, t is the current time, Δt is the time elapsed, α ($0<\alpha<1$) is the smoothing constant, S^{se}_t is the first-order exponential smoothing prediction model, and S^{de}_t is the double-smoothed statistics.

$$\begin{aligned} P_{t+\Delta t} &= (2+\alpha\cdot\Delta t/(1-\alpha))\cdot S_t^{se} - (1+\alpha\cdot\Delta t/(1-\alpha))\cdot S_t^{de} \\ S_t^{se} &= \alpha\cdot M_t + (1-\alpha)\cdot S_{t-1}^{se} \\ S_t^{de} &= \alpha\cdot S_t^{se} + (1-\alpha)\cdot S_{t-1}^{de} \end{aligned} \tag{4}$$

We adopt a fuzzy logic control system to the proposed agent for autonomic and dynamic load balancing. The predicted data volume is used as fuzzy input parameters; the input X is the volume of tag data and the input Y is the volume of cache data. And, the output Z is the workload of an edge M/W. Each fuzzy parameter contains fives fuzzy sets. As shown in Fig. 1, the fuzzy sets are mapped to membership functions with values between 0 and 1. Inputted data volume is converted from crisp data to fuzzy data according to its own membership function. The range of tag data and cache data fuzzy sets represent linguistic variables such as Very-Small, Small, Normal, Large, and Very-Large. Also, the range of work load fuzzy sets represents linguistic variables such as Very-Low, Low, Medium, High, and Very-High.

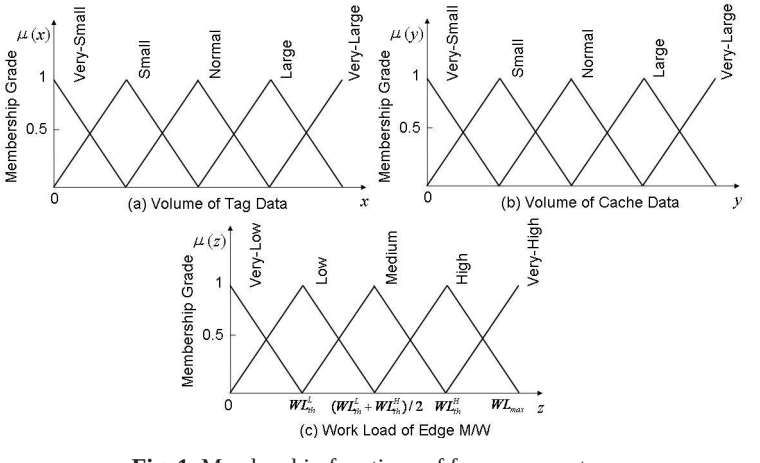


Fig. 1. Membership functions of fuzzy parameters

Membership functions of fuzzy sets are combined with relevant fuzzy control rules to begin the fuzzy inference process. In this paper, a total of 45 fuzzy control rules are adopted as shown in Table 1. These rules are directly related to the workload status of an edge M/W and can be represented by the IF-THEN rule. For example, R12 in Table 1 can be expressed as: (R12) IF the volume of tag data is Small and the volume of cache data is Normal, TEHN work load of an edge M/W is Low.

Table 1. Fuzzy control rules (Rows and columns represent two inputs; the volume of tag data (X) and the volume of cache data (Y). These inputs refer IF parts in IF-THEN rules. And, the intersection of each row and each column refers to THEN part in IF-THEN rules.)

Y \ X	Very Small	Small	Normal	Large	Very Large
Very Small	(R1) Very Low	(R2) Very Low	(R3) Low	(R4) Low	(R5) Normal
Small	(R6) Very Low	(R7) Low	(R8) Low	(R9) Normal	(R10) High
Normal	(R11) Low	(R12) Low	(R13) Normal	(R14) High	(R15) High
Large	(R16) Low	(R17) Normal	(R18) High	(R19) High	(R20) Very High
Very Large	(R21) Normal	(R22) High	(R23) High	(R24) Very High	(R25) Very High

Also, we adopt the max-min method of Mamdani [10] for fuzzy inference. The work load state of an edge M/W is estimated from a result of fuzzy inference. And, defuzzification is needed to convert the work load state, which is expressed in terms of fuzzy sets, into a real value, which indicates the volume of work load. We use the center of gravity (COG) method [10] for defuzzification. Therefore, the real value is calculated as the equation (5), where WL_{vol} is the volume of work load, c is the compatibility with the fore part of a fuzzy rule and $S_{\max}$ is the maximum singleton value of an output fuzzy set, and WL_{max} is the volume of maximum tag and cache data.

$$WL_{vol} = (\sum_{i=1}^{n} (S_{\max i} \times c_i) / \sum_{i=1}^{n} c_i) \times WL_{\max} \tag{5}$$

3.2 Dynamic RFID Load Balancing Algorithm

Fig. 2 shows pseudo code of the proposed algorithm. The fuzzy logic control based-load balancing agent executes the load balancing algorithm in comparison the workload volume $WL(em)$ of an edge M/W with two thresholds; WL^H_{th} and WL^L_{th}. $WL(em)$ is an estimated fuzzy output value described in the previous section. WL^H_{th} and WL^L_{th} indicate the maximum and the minimum workload thresholds which are allowed to process without data loss and delay. These thresholds are settled by CPU speed, memory capacity, and queue size of a host computer equipped with an edge M/W. Before the load balancing time, we sort a reader list by workload in an ascending order to decide how much workload to immigrate into other edge M/Ws. The reader

list contains the volume of tag and cache data by each reader. As shown in Fig. 2, we divide the workload volume of an edge M/W into three states: '*Over Loaded*', '*Under Loaded*', and '*Well Loaded*', explained as follows:

Over Loaded ($WL(em) > WL^{H}_{th}$) – In this situation, the agent calculates an overload, which indicates a difference between the evaluated workload volume and the maximum workload threshold. The current workload is reduced as much as the overload. Also, the agent searches a target edge M/W, which has a surplus capacity suitable to process the overload, from edge M/W information. And, the overload is immigrated from the current edge M/W to the target edge M/W.

Under Loaded ($WL(em) < WL^{L}_{th}$) – In this case, we consider that the current workload is inferior to computing performance of an edge M/W. Therefore, the agent requests a target edge M/W, which has the closest overload to the surplus capacity of an edge M/W managed by the agent, to transfer the overload to itself.

Well Loaded ($WL^{L}_{th} \leq WL(em) \leq WL^{H}_{th}$) – The current workload of an edge M/W with the '*Well Loaded*' state is adequate for the computing ability of the edge M/W. Therefore, the agent does not execute RFID load balancing and only registers its workload information to an integration M/W.

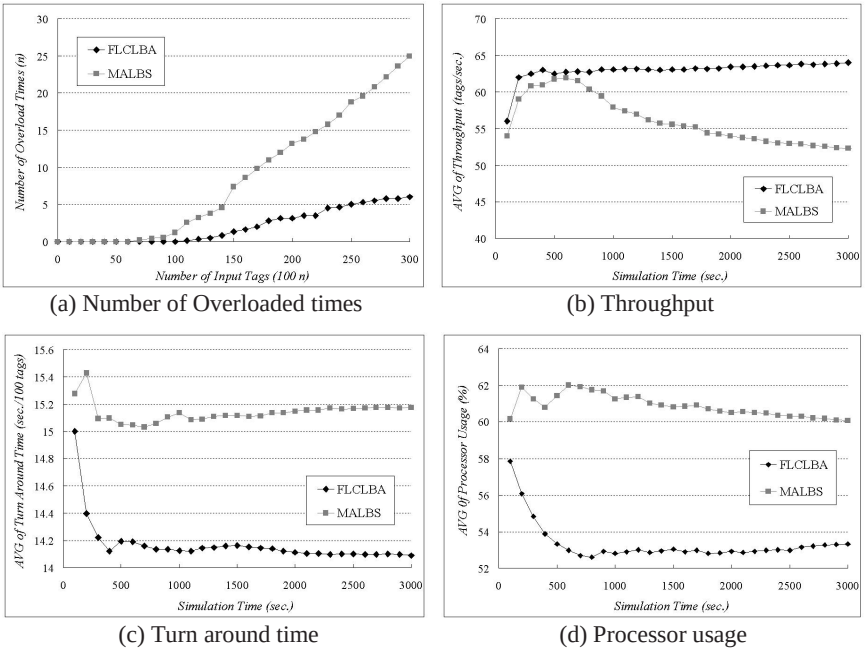
```
// sorting readers in increasing order of work load
for i = 1 to m-1
    if WL(ReaderList[i]) > WL(ReaderList[i+1])
    swap ReaderList[i] with ReaderList[i+1]
    end if
    increase i
end for
get Edge M/W information from Integration M/W
// case of over-loaded edge M/W
if WL(em_a) > WL^H_th(em_a)
   // calculating overload of Edge M/W
   j = size of ReaderList
   while (OL(em_a) < (WL(em_a) – WL^H_th(em_a)))
      OL(em_a) + = WL(ReaderList[j])
      WL(em_a) – = WL(ReaderList[j])
      decrease j
   end while
   // selecting Edge M/W to reallocate overload
   for k = 1 to n-1
      if WL(em_k) + OL(em_a) ≤ WL^H_th(em_k)
         E1(em_k) = (UL(em_k) – OL(em_a)) / UL(em_k)
      else E1(em_k) = 1
      end if
      increase k
   end for
   target1 = Edge M/W with Min(E1(em_k))
   reallocate OL(em_a) to target
end if
// case of under-loaded edge M/W
if WL(em_a) < WL^L_th(em_a)
   // calculating underload of edge M/W
   UL(em_a) = WL^H_th(em_a) – WL(em_a)
   // selecting Edge M/W to request workload
   for l = 1 to n-1
      if WL(em_a) + OL(em_l) ≤ WL^H_th(em_a)
         E2(em_l) = (OL(em_l) –UL(em_a)) / OL(em_l)
      else E2(em_l) = 1
      end if
      increase l
   end for
   target2 = Edge M/W with Min(E2(em_l))
   request target2 to OL(em_l)
end if
// case of well-loaded edge M/W
if WL^L_th(em_a) ≤ WL(em_a) ≤ WL^H_th(em_a)
   OL(em_a) = 0
   UL(em_a) = WL^H_th(em_a) – WL(em_a)
   register UL(em_a) to Integration M/W
end if
```

Fig. 2. Dynamic RFID load balancing algorithm

4 Experimental Results

In order to evaluate efficiency and effectiveness of the fuzzy logic control-based load balancing agent (FLCLBA), we implemented its prototype on the DEVS modeling and simulation environment [8]. And, we compared system performance of the

FLCLBA with that of the mobile agent-based load balancing system (MALBS) [7] of which a load balancing agent relocates readers of an overloaded edge M/W to other edge M/Ws directly. In experiments, we used six edge M/Ws and one integration M/W to build a distributed RFID system. And, a reader emulator was used in generating and capturing tag readings. Also, high and low thresholds for RFID load balancing were fixed to 80% and 20% of edge M/W's workload.



(a) Number of Overloaded times (b) Throughput

(c) Turn around time (d) Processor usage

Fig. 3. Comparison of Performance Measures (Fuzzy logic Control-based Load Balancing Agent (FLCLBA) vs. Mobile Agent-based Load Balancing System (MALBS))

Fig. 3(a) shows variations of the number of overload times by the number of input tags. The number of overload times indicates how many times edge M/Ws run into the '*Over Loaded*' state. The FLCBA provides about 75% reduction in overload times as compared with the MALBS. When 30000 tags were processed on a RFID system, in case of the FLCLBA, the number of overload times totaled 6, while the MALBS recorded a total of 25 overload times. Fig. 3(b) shows variations of throughput by simulation times. Throughput is the number of processed tags on an edge M/W per one second. The average throughput of the FLCLBA is 63.11 tags and that of the MALBS is 55.52 tags. The FLCLBA shows about 13.7% increment in throughput as against the MALBS. Also, as shown in Fig. 3(b), the FLCLBA gradually increased in throughput, while the MALBS decreased. Because, the MALBS generated more overloaded edge M/Ws than the FLCLBA, which causes tag loss and delay. Fig. 3(c) shows variations of turn around time by simulation times. Turn around time indicates the elapsed time from generation of a tag to processing completion of the tag. On the average, the turn around time of the FLCLBA was 14.16 seconds, which is about

6.5% less than that of the MALBS. During simulations, the MALBS dissipated communication times to relocate readers. This increase of communication times resulted in an increased latency for data processing. Fig. 3(d) shows variations of processor usage by simulation times. On the average, the FLCLBA recorded 53 % processor usage to process generated tags, while the MALBS recorded about 61% processor usage. The FLCLBA provides about 12.2% reduction in processor usage as compared with the MALBS. These results demonstrate that the proposed agent is more effective to load balancing of a distributed RFID system than the MALBS.

5 Conclusion

In this paper, we have proposed the fuzzy logic control-based load balancing agent for distributed RFID systems. The proposed agent predicts data volume of an edge M/W and evaluates workload of the edge M/W by fuzzy logic control. In order to maintain stable workload, the agent immigrates overload into other under-loaded edge M/Ws or requests data reallocation to other overloaded edge M/Ws according to the dynamic RFID load balancing algorithm. Experimental results show that the proposed agent enhances system availability and scalability with quick and steady RFID data processing in comparison with the mobile agent-based load balancing system [7].

References

1. Brown, D.: Rfid Implementation. McGraw-Hill, New York (2006)
2. Derakhshan, R., Orlowska, M., Li, X.: RFID Data Management: Challenges and Opportunities. In: IEEE International Conference on RFID 2007, pp. 175–182. IEEE Press, Los Alamitos (2007)
3. EPCglobal Inc., http://www.epcglobalinc.com
4. Park, J.G., Chae, H.S., So, E.S.: A Dynamic Load Balancing Approach Based on the Standard RFID Middleware Architecture. In: IEEE International Conference on e-Business Engineering 2007, pp. 337–340. IEEE Computer Society, Los Alamitos (2007)
5. Park, S.M., Song, J.H., Kim, C.S., Kim, J.J.: Load Balancing Method Using Connection Pool in RFID Middleware. In: The 5th ACIS International Conference on Software Engineering Research, Management & Applications, pp. 132–137. IEEE Computer Society, Los Alamitos (2007)
6. Dong, Q., Shukla, A., Shrivastava, V., Agrawal, D., Banerjee, S.: Load Balancing in Large-Scale RFID Systems. In: The 26th IEEE International Conference on Computer Communication, pp. 2281–2285. IEEE Press, Los Alamitos (2007)
7. Cui, J.F., Chae, H.S.: Developing Load Balancing System for RFID Middlewares Using Mobile Agent Technology. In: Nguyen, N.T., Grzech, A., Howlett, R.J., Jain, L.C. (eds.) KES-AMSTA 2007. LNCS (LNAI), vol. 4496, pp. 757–764. Springer, Heidelberg (2007)
8. Zeigler, B.P., Sarjoughian, H.S.: Distributed simulation: creating distributed simulation using DEVS M&S environments. In: Winter Simulation Conference 2000, pp. 158–160 (2000)
9. Bowerman, B., O'Connell, R., Koehler, A.: Forecasting, Time Series and Regression, 4th edn. Thomson Books/Cole (2004)
10. Bai, Y., Zhuang, H., Wang, D.: Advanced Fuzzy Logic Technologies in Industrial Applications. Springer, London (2006)

A Learning Assistance Tool for Enhancing ICT Application Ability of Elementary and Secondary School Students

Chenn-Jung Huang, Yun-Cheng Luo, Chun-Hua Chen, Tun-Yu Chang, Kai-Wen Hu, Tsung-Hsien Wu, and Kuo-Liang Huang

Department of Computer Science and Information Engineering, National Dong Hwa University
cjhuang@mail.nhlue.edu.tw

Abstract. In recent years, developing useful learning assistance systems has become a hot research topic in the literature. The learner can be benefited by the useful guidance provided by the learning assistance tool. An effective learning assistance tool can reduce the teaching load of the teacher. However, it is rarely seen that a learning assistance tool employs machine learning techniques to provide appropriate diagnosis or feedback to learners or teachers in the literature. We thus propose a learning assistance tool that employs reinforcement learning technique to continuously interact with the environment in order to offer learners suitable and timely feedback, guide them through the difficulties. Our experimental results reveal that our learning assistance tool can effectively enhance the learners' ICT application ability and assist the learners in overcoming difficulties. The teaching load of the teacher is also significantly reduced.

Keywords: Information and communication technology, application software, learning assistant, diagnosis, machine learning, reinforcement learning.

1 Introduction

Numerous web-based E-learning platforms and examination systems have been extensively developed in the literature. These ICT-based systems not only reduce the efforts needed in grading, marking, recording, analyzing and other time-consuming examination-related operations, but also allow practices and tests to be attempted anywhere and anytime through the use of web browsers [1],[8],[10].

Observing that students can learn more effectively and be encouraged if there exists a learning assistance tool providing useful hints or feedbacks to them, we thereby develop a learning assistance tool to give the learners the guidance when the learner is confused or stalled in using an ICT Application ability assessment platform that we built for elementary and secondary school students in Taiwan earlier [5]. Project based activities are incorporated into the web-based diagnosis and assessment platform and the students are expected to apply application software suite provided by the platform to solve the daily life problem. The guidance is offered via a feedback rule construction mechanism wherein a machine learning technique, reinforcement learning, is adopted to provide

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 661 – 668, 2008.
© Springer-Verlag Berlin Heidelberg 2008

useful hints or feedbacks to learners based on their learning portfolios collected during their learning activities.

The remainder of the paper is organized as follows. Section 2 gives the overview of the related work. Section 3 describes the overall architecture of the proposed ICT learning assistance tool. Section 4 gives the experimental results. Conclusions and the future work are given in Section 5.

2 Relate Work

In the past few years, developing useful learning assistance systems has become a hot research topic in the literature. A web-based mathematics learning assistance system (Melis *et al.* 2007) [9], ActiveMath, was developed to assist learners in searching the interesting courses or practical examples for their study. Chang (2005) [3] developed an on-line learning diagnosis system to assist teacher in monitoring learners' learning status in time.

All of the above-mentioned learning assistance systems more or less employed the machine learning techniques to provide appropriate assistance to learners or teachers. We thus attempt to adopt a well-known machine learning technique, reinforcement learning, to keep track of the learners' operations and provide some suggestions to the learners when the learners are confused or stalled in using the ICT Application ability assessment platform we developed earlier.

Reinforcement learning algorithms have been developed that are closely related to methods of dynamic programming, which is a general approach to optimal control [2],[4],[6], [7]. To the best of our knowledge, reinforcement learning technique has never been applied to the design of intelligent e-learning platform in the literature.

3 Architecture of Learning Assistance Tool

Fig.1. shows the architecture of the learning assistance tool used in this work. There are four main components in the e-learning platform, including a script content/learning activity management module, the learners' workspace interface, a learning diagnosis module, and an assessment module.

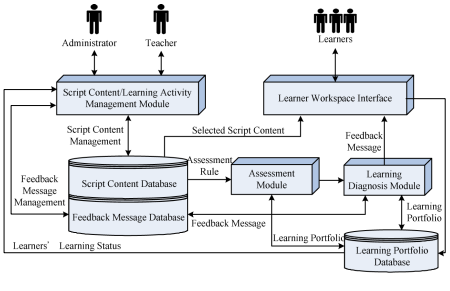


Fig. 1. Architecture of learning assistance tool

3.1 Script Content/Learning Activity Management Module

The teacher who is willing to adopt the learning assistant tool in her/his class can choose some appropriate script from the script content database, and can setup the learning activity time period and specify the class that attends the learning activity as illustrated in Fig. 2.

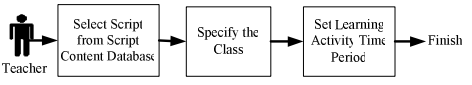


Fig. 2. Process of learning activity setting

3.2 Students' Workspace Interface

Five kinds of application software modules are provided in this work. Three open source software modules are modified to mimic the functions provided by Microsoft application suite, Word, Excel and PowerPoint. In addition, two more tools including built-in search engine and e-mail management are also developed to fit the project.

3.3 The Assessment Module

The automatic assessment module includes two main components, expected result generator and comparison unit. The expected result generator assists the teacher who designed the project to generate the expected result after each step sequence. The expected results are saved in the database for comparing each student's answers.

3.4 The Learning Diagnosis Module

The learning diagnosis module can provide reference for teachers or our system to give necessary assistance for students when the student encounters a problem at a specific step. As the reward value is lower than some preset threshold, the learning assistance mechanism would give an appropriate feedback based on the corresponding environment and the reward setting.

4 Application of Reinforcement Learning Mechanism to Learning Diagnosis Module

Figure 3 shows the reinforcement learning's normal operation. An agent learns effective decision-making policies through an online trial-and-error process in which it interacts with an environment. Each interaction consists of

- observing the environment's current state s_t at time t,
- performing some legal action at in state s_t, and
- receiving a reward r_t, which is a numerical value that the user would like to maximize, followed by an observed transition to a new state s_{t+1}.

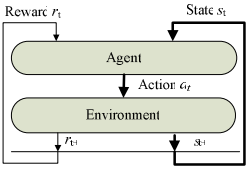


Fig. 3. Reinforcement learning mechanism

The *Q-learning* and *Temporal Differences* approaches of reinforcement learning algorithm are most likely to be used by other researchers. The general value function is as follows,

$$R = \sum_{k=1}^{n} \gamma^k \times r_k \tag{1}$$

where R is general reinforcement learning value function. It sets every path probability which learners choose to be γ, and reward r *is* multiplied by γ. k is the step number of the path probability.

The probability of user choosing the action from j actions in ith state

$$p_{i,j} = \frac{n_j}{n_i} \tag{2}$$

where n_i is the total number of users at ith state and n_j is the number of users choosing jth action.

A threshold is set to determine if the learner can finish the action within an expected time interval. If the learner has not finished this mission within the expected time interval, a time reward will be sent to the learner as follows,

$$t_{reward} = \begin{cases} 0 & , \text{if } t < t_{thresshold} \\ t \times r & , \text{if } t > t_{thresshold} \end{cases} \tag{3}$$

where $t_{thresshold}$ is a preset timeout, t denotes the elapse time after timeout, and r is a time reward which is given by experts.

The total reward is defined as the selection of the best way to proceed with. It is expressed by,

$$R = (\sum_{i=1}^{n} (\prod_{j=1}^{m} \max(p_{i,j})) \times reward_{i,j}) - t_{reward} \tag{4}$$

where max($P_{i,j}$) represents the most possible action, which is indexed by j, selected during the ith stage. $reward_{i,j}$ is the reward which is given by experts for each path.

Fig. 4. An example of weak hint

Notably, the reward, t_{reward}, is subtracted per unit time in this equation because the probability of completing the mission keeps decreased when the clock is ticking. In our platform, there are there levels of hints including weak, moderate, and strong hints. Figure 4 shows an example of the weak hint message.

5 Experimental Result and Analysis

The experimental group and the control group consist of 29 fifth-grade pupils each. The experimental group and the control group are both experimented with three tests. In the control group, no learning 130 assistance mechanism was provided during the assessment. They should try to find useful information in the assessment platform by themselves to complete the mission. On the other hand, the experimental group is aided by the learning assistance tool so that appropriate hints or feedbacks can be prompted to the learners timely. Post-test was given to the experimental group and the control group by using the assessment platform without the usage of the learning assistance tool.

Table 1 and Table 2 show the T test result of the pre-test. There is no obvious difference between the experimental group and the control group. It can be inferred that the students in both groups possess similar basic computer operation skills. Table 3 to The correlation between pre-test achievement and post-test achievement for the experimental group and the control group using Pearson correlation coefficient analysis was illustrated in Table 8 show the T test results of post-test between the high, medium, and low achievement groups. It is observed from Table 3 and Table 4 that there is an obvious difference between the experimental group and the control group in the high achievement groups, whereas no significant difference between the experimental group and the control group exists in the medium and low achievement groups.

Table 1. Group statistics of pre-test

Group	N	Mean	Std. Deviation	Std. Error Mean
The experimental group	29	78.966	16.4414	3.0531
The control group	29	79.655	18.4709	3.2443

Table 2. Independent sample T test of pre-test

	Levene's test for equality of variances		T-test for equality of means					
							95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
With qual variance	0.253	0.617	-0.15	56	0.878	-0.6897	-9.614	8.2347
Without equal variance			-0.15	55.79	0.878	-0.6897	-9.614	8.2354

Table 3. Group statistics of post-test for high achievement group

Group	N	Mean	Std. Deviation	Std. Error Mean
The experimental group	10	82.00	20.440	6.464
The control group	10	59.00	26.854	8.492

Table 4. Independent sample T test of post-test for high achievement group

	Levene's test for equality of variances		T-test for equality of means					
							95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
With equal variance	0.858	0.367	2.155	18	0.045	23.000	0.579	45.421
Without equal variance			2.155	16.808	0.046	23.000	0.465	45.535

Table 5. Group statistics of post-test for medium achievement group

Group	N	Mean	Std. Deviation	Std. Error Mean
The experimental group	10	59.00	24.244	7.667
The control group	10	51.00	15.239	4.819

Table 6. Independent sample T test of post-test for medium achievement group

	Levene's test for equality of variances		T-test for equality of means					
							95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
With equal variance	5.038	0.038	0.883	18	0.389	8.000	-11.02	27.025
Without equal variance			0.883	15.151	0.391	8.000	-11.28	27.284

Table 7. Group statistics of post-test for low achievement group

Group	N	Mean	Std. Deviation	Std. Error Mean
The experimental group	9	36.67	20.616	6.872
The control group	9	41.11	20.883	6.961

The correlation between pre-test achievement and post-test achievement for the experimental group and the control group using Pearson correlation coefficient analysis was illustrated in Table 9 and Table 10 shows that a significant positive correlation

Table 8. Independent sample T test of post-test for low achievement group

	Levene's test for equality of variances		T-test for equality of means					
							95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
With equal variance	0.063	0.806	-0.454	16	0.656	-4.444	-25.180	16.292
Without equal variance			-0.454	15. 997	0.656	-4.444	-25.181	16.292

(Sig.=0, r=0.66) was found between the pre-test and post-test achievements in the experimental group, whereas medium positive correlation (Sig.=0, r=0.66) was observed in the control group as show in Table 10.

Table 9. Pearson correlation coefficient analysis in the experimental group

		Pre-test	Post-test
Pre-test	Pearson Correlation	1	.660**
	Sig. (2-tailed)		.000
	N	29	29
Post-test	Pearson Correlation	.660**	1
	Sig. (2-tailed)	.000	
	N	29	29

** Correlation is significant at the 0.01 level

Table 10. Pearson correlation coefficient analysis in the control group

		Pre-test	Post-test
Pre-test	Pearson Correlation	1	.423*
	Sig. (2-tailed)		.022
	N	29	29
Post-test	Pearson Correlation	.423*	1
	Sig. (2-tailed)	.022	
	N	29	29

*Correlation is significant at the 0.05 level

6 Conclusions and Future Work

The experimental results reveal that the students will achieve better performance after using learning assistance system in the experimental group. The learning assistance tool proposed in this work is thus proved to be helpful for students. The teaching load

of the teacher is also significantly reduced because the appropriate hints or feedbacks can be automatically prompted to the learners without the involvement of the teachers. In the future work, we will utilize students' learning portfolio and adopt appropriate machine learning techniques to establish the probability model of each action node on the traversing path toward completion of the assigned task and automatically determine the reward value at each node that is currently set by the domain experts.

Acknowledgements. This research was partially supported by Ministry of Education and National Science Council of Taiwan (Contract No. NSC 96-2628-E-026-001-MY3).

References

1. Barsky, D.B., Shafarenko, A.V.: WWW and Java-Based Distributed Examination System for Distance Learning Applications. In: Second Aizu International Symposium on Parallel Algorithms/Architecture Synthesis, pp. 356–363 (1997)
2. Blumberg, B., Downie, M., Ivanov, Y., Berlin, M., Johnson, M., Tomlinson, B.: Integrated Learning for Interactive Synthetic Characters. In: Proceedings of the ACM SIGGRAPH (2002)
3. Chang, G.E.: Promoting The Information Talent: Define Information Competence Indicators of High School (2005)
4. Evans, R.: Varieties of Learning. In: Rabin, S. (ed.) AI Game Programming Wisdom, pp. 567–578. Charles River Media, Hingham (2002)
5. Huang, C.J., Liu, M.C., Tsai, C.H., Chang, T.Y., Luo, Y.C., Chen, C.H., Chen, H.X.: Developing Web-Based Information Competence Assessment Platform for Grade 1-9 Curriculum in Taiwan. In: Proceedings of 2007 Taiwan Academic Network Conference, pp. 1–9 (2007)
6. Isbell, C., Shelton, C., Kearns, M., Singh, S., Stone, P.: Cobot: A Social Reinforcement Learning Agent. In: 5th Intern. Conf. on Autonomous Agents (2001)
7. Kaplan, F., Oudeyer, P.Y., Kubinyi, E., Miklosi, A.: Robotic Clicker Training. Robotics and Autonomous Systems 38(3–4), 197–206 (2002)
8. McGough, J., Mortensen, J., Johnson, J., Fadali, S.: A Web-Based Testing System with Dynamic Question Generation. In: 31st ASEE/IEEE Frontiers in Education Conference, Reno, NV, vol. 3, S3C-23-8, pp. 3–23 (2001)
9. Melis, E., Ullrich, C., Goguadze, G., Libbrecht, P.: How ActiveMath Supports Moderate Constructivist Mathematics Teaching. In: 8th International Conference on Technology in Mathematics Teaching (2007)
10. Shafarenko, A., Barsky, D.: A Secure Examination System with Multi-Mode Input on the World-Wide Web. In: International Workshop on Advanced Learning Technologies, Palmerston North, NZ, pp. 97–100 (2000)

A Sliding Singular Spectrum Entropy Method and Its Application to Gear Fault Diagnosis

Yong Lu, Yourong Li, Han Xiao, and Zhigang Wang

College of Mechanical Engineering and Automation, Wuhan University of Science and Technology, Wuhan 430081, P.R. China
{lvyong,liyourong,xiaohan,wangzhigang}@wust.edu.cn

Abstract. Entropy changes with the variation of the system status. It has been widely used as a standard for the determination of system status, quantity of system complexity and system classification. Based on the singular spectrum entropy of traditional calculation method, a sliding singular spectrum entropy method is proposed to use for singularity detection and extraction of impaction signal. Each original signal point is intercepted a neighborhood points of the signal with a given length and the singular spectrum entropy for the intercepted signal is calculated. A surrogate signal with the same length as the original signal is acquired by point-to-point calculation. Numerical simulation and gear fault diagnosis experiment are studied to verify the proposed method, the results show that the method is valid for the reflection on the changing of system status, singularity detection and the extraction of the weak fault feature signal mixed in the strong background signal.

Keywords: Nonlinear time series analysis; singular spectrum entropy; fault diagnosis.

1 Introduction

The analysis of data such as machinery vibration signal collected in industry is an important work. Many techniques and tools have been developed and implemented. Entropy is a useful tool for investigating the dynamics of signals. It can be used to measure the complexity of various signals and quantify information in signals such as electrocardiograph, machinery vibration signal and others. Entropy supplies topological information about the folding process, which can be used as a classification tool. Grassberger (1987) [1] has proposed to calculate the chaos characteristics of system using the phase space reconstruction such as fractal dimension, Lyapunov exponent and Komologrov entropy, etc. Phase space reconstruction has found wide applications in different fields since it was proposed. After the time series were reconstructed from one dimension to higher dimension, it indicates that the implicit structure feature of nonlinear phenomenon and its dynamic evolution rule can be acquired by the geometrical property of the phase point in high dimension phase space [2]. LU Zhiming(2000)[3] adopts singular value decomposition to the attractor after phase space reconstruction, the singular spectrum of high dimension data is acquired, several maximum singular value is kept, and the others are set to zero, then the signal is

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 669–678, 2008.
© Springer-Verlag Berlin Heidelberg 2008

reconstructed using inverse transformation and the noise reduction is realized by the transformation. YANG Wenxian (2000) [4] has studied the characteristic of the singular spectrum entropy for the machinery vibration signal after phase space reconstruction. Stephen J. Roberts [5] has proposed a segmentation singular spectrum entropy method, the temporal and spatial complexity measures for EEG signal is analyzed with it. The singular spectrum entropy in traditional research is used as a single quantification target mentioned earlier, in fault diagnosis, entropy is used for fault classification as an identification parameter [6]. Based on the studies of Roberts S J[7], the application of entropy is extended to the fault signal feature extraction in this paper, the dynamic changing signal is redescribed adopting entropy, a sliding singular spectrum entropy is proposed, from which hidden impact feature is gained. In order to acquire the accurate position of the hidden structure in original signal, the length of the sliding singular spectrum entropy is the same as the original signal. Compared with the methods from other researchers, the sliding singular spectrum entropy is a new description method of original signal, which describe the hidden signal feature from a new point of view. The structure of modern machinery is more and more complex; the vibration signal of broken machinery we collect is usually nonlinear and non-stationary, especially when the machinery has the early fault, the signal that comes from the broken part is so weak that we can't identify it from the strong background signal. The aim of the fault signal feature extraction is to gain the information which can express the fault feature from the samples. The sliding singular spectrum entropy of original signal is calculated, and then the hidden fault feature in the original signal can be acquired from it. A new valid method for the machinery online monitoring and fault diagnosis is given in the paper.

2 The Calculation Method of Sliding Singular Spectrum Entropy

The reconstruction of a time series which is equivalent to the original state space of a system form a scalar time series is almost the basis of all nonlinear methods. Takens' embedding theorem [2] has proved that it is feasible. For one dimension time series, the m dimension phase space with the same dynamic characteristic as the original system can be gained by reconstruction.

For a given one dimension time series $\{x_i\}$ $(i=1,2,\cdots,N)$, the high dimension state vector space after reconstruction can be expressed as

$$X_i=[x_i, x_{i+\tau},\ldots, x_{i+(m-1)\tau}]. \tag{1}$$

X_i is the ith phase point, which expresses a state of m dimension phase space. m is the embedding dimension and τ is the delay time.

The reconstruction phase space can be expressed as the equation (1), which makes the original time series reconstructed from one dimension to higher dimension. We can get a reconstructed vector Y by the equation (1).

$$Y=[X_1, X_2,\cdots, X_{N-(m-1)\tau}]^T. \tag{2}$$

The singular value decomposition of the reconstructed vector Y can be calculated. For a given real vector matrix $Y_{m\times n}$, if the rank of it is r, there are two standard orthogonal matrix $U_{m\times m}$, $V_{n\times n}$ and a diagonal matrix $D_{m\times n}$, which makes the equation (3) come into existence.

$$Y=UDV^T , \tag{3}$$

where $D_{m\times n}=\begin{bmatrix}\Sigma_{r\times r} & 0\\ 0 & 0\end{bmatrix}$, $\Sigma_{r\times r}=diag(\sigma_1,\sigma_2,...,\sigma_r)$ $U_{m\times m}=(u_1,u_2,...,u_r,u_{r+1},...,u_m)$ $V_{n\times n}=(v_1,v_2,...,v_r,v_{r+1},...,v_n)$

And $\sigma_i=\sqrt{\lambda_i}$ $(i=1,2,...,r,...,n)$ can be called the singular value of the matrix Y, where $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_r\geq 0$, $\lambda_{r+1}=\lambda_{r+2}=\cdots=\lambda_n=0$ are the eigenvalues of matrix Y^TY and YY^T, and u_i,v_i $(i=1,2,...,r)$ are the eigen vectors corresponding the non-zero eigenvalues λ_i of matrix Y^TY and YY^T.

Supposing the singular value after decomposition is σ_i , and defining p_i as the equation (4).

$$p_i=\frac{\sigma_i}{\sum_{i=1}^{m}\sigma_i} . \tag{4}$$

The singular spectrum entropy of a signal can be defined as equation (5).

$$H=\sum_{i=1}^{m}p_i\log p_i \tag{5}$$

The system complexity can be defined based on the equation (5), in the equation, if we choose 2 to be the logarithmic base and define $A(H)$ as system complexity [7], then we can write the system complexity $A(H)$ as the equation (6).

$$A(H)=2^H . \tag{6}$$

Stephen J. Roberts [7] have used equation (6) for monitoring the temporal and spatial complexity of the data, based on the above segmentation thoughts, we use the system complexity definition （6） as a signal feature extraction method in this paper.

The procedure for applying the proposed sliding singular spectrum entropy for signal feature extraction method is detailed below.

Step 1. Acquiring the sliding calculation point of the original signal $\{x_i\}$ $(i=1,2,\cdots,N)$ with length N , i.e. for each original point, we constitute a time series with length k . In order to make the processing time series have the same length with the original time series, the original time series $\{x_i\}$ is added with a time series with length $k/2$ in the first point and the last point, the front time series with length

$k/2$ have the same value that equals to the average of the first k points of original time series, and the back time series with length $k/2$ have the same value that equals to the average of the last k points of original time series.

Step 2. Each original point is processed according to the step 1), because the length of original time series equals to N, we can get N groups time series with each group has k points, and then the linear tendency component of the N groups is removed.

Step 3. Each group is calculated by the equation (1) ~ (5), and the N groups singular spectrum entropy values $A(H)$ can be gained. The new time series $\{A_i\}$ $(i=1,2,\cdots,N)$ with the same length as the original time series is defined as the sliding singular spectrum entropy time series $A(t)$.

With this approach, impulsive features are enhanced while the background noises are suppressed. The time series adopting sliding singular spectrum entropy method by the step 1 ~ 3 can be used to express the hidden non-stationary transformation feature that we can't simply gain from the original time series.

The simulation experiment using the sliding singular spectrum entropy method is studied in the paper, the length k is given as 20, the dimension m is given as 10 and the delay time τ is given as 1.

3 Signal Feature Extraction Study Using the Sliding Singular Spectrum Entropy Method

In order to test the validity of the sliding singular spectrum entropy method, the numerical value simulation is given on the following.

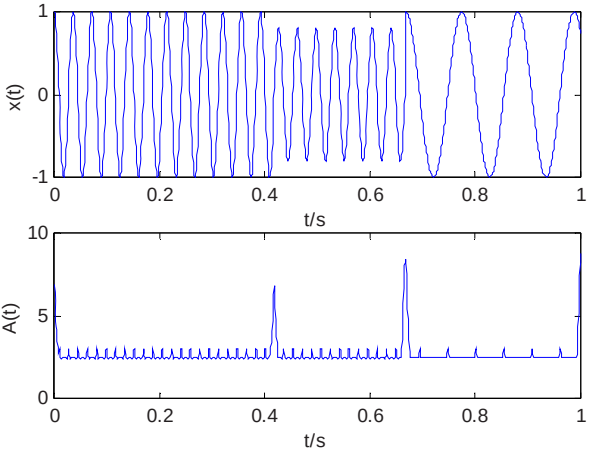
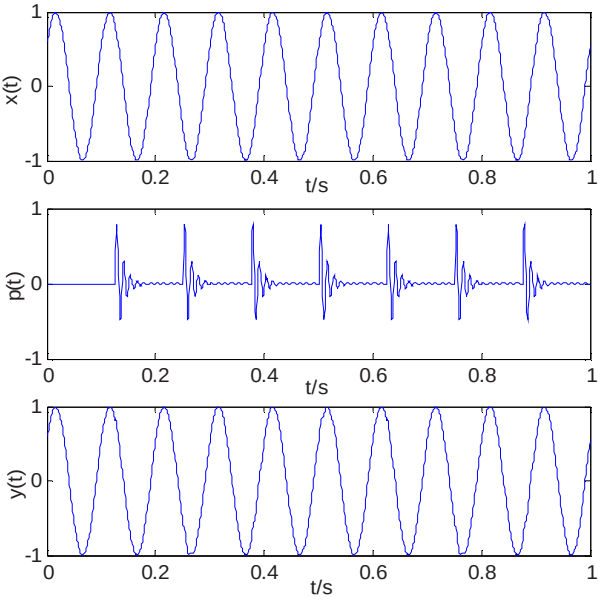


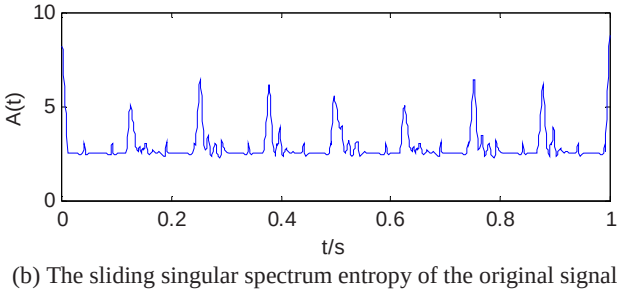
Fig. 1. Signal singularity detection using sliding singular spectrum entropy method

3.1 Signal Singularity Detection Using the Sliding Singular Spectrum Entropy Method

Frequency changing and amplitude changing signal is processed with the sliding singular spectrum entropy method. The original signal is composed by frequency changing and amplitude changing components in turn, the original signal is displayed as the $x(t)$ in figure 1, the sliding singular spectrum entropy of the signal is displayed as $A(t)$ in figure 1. There are several obvious changing positions in the sliding singular spectrum entropy $A(t)$, which are consistent with the beginning changing position of the signal with the different frequency and amplitude.



(a) Original signal $x(t)$, $p(t)$ and $y(t)$



(b) The sliding singular spectrum entropy of the original signal

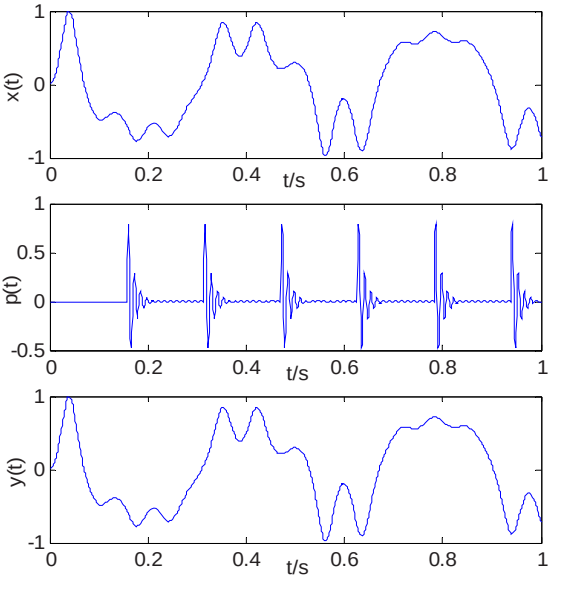
Fig. 2. The extraction of weak impact signal mixed into the sine signal

3.2 Hidden Impact Signal Extraction Using Sliding Singular Spectrum Entropy Method

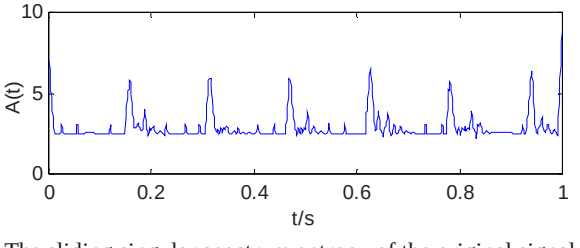
According to above simulation experiment results, the sliding singular spectrum entropy will change with the signal state. Besides, we can abstract the periodical impact signal hidden in the strong background signal based on the before-mentioned idea. The simulation experiment on abstracting weak impact signal mixed into the sine and chaos signal is given to validate the method.

3.2.1 The Extraction of Weak Impact Signal Mixed into the Sine Signal

Supposing the sine signal is $x(t) = \sin(20\pi t + 0.2\pi)$, the simulation impact signal is $p(t) = 0.1 \cdot \exp(-5t)\sin(10\pi t)$, the two original signals and the added signal $y(t)$ are displayed on the figure 2(a). The sliding singular spectrum entropy $A(t)$ of signal $y(t)$ is given on the figure 2(b).



(a) Original signal $x(t)$, $p(t)$ and $y(t)$



(b) The sliding singular spectrum entropy of the original signal

Fig. 3. The extraction of weak impact signal mixed into the chaos signal

It is difficult to make out the impact signal from the original signal $y(t)$ displayed on the figure 2(a), but we can obviously make out the impact signal from the processed signal $A(t)$ displayed on the figure 2(b).

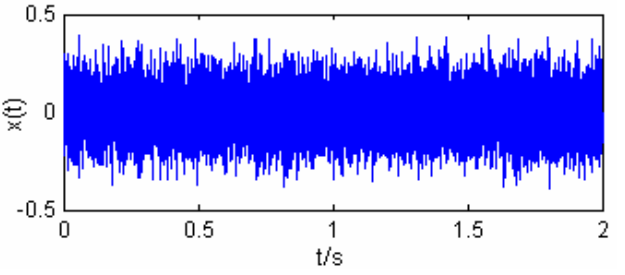
3.2.2 The Extraction of Weak Impact Signal Mixed into the Chaos Signal

By using Runge-Kutta integration method, the chaos signal is produced by the Duffing differential equation $\ddot{x}+0.2\dot{x}-x+x^3=40\cos t$, the chaos signal is normalized, and then, it is added with the impact signal $p(t)=0.1\cdot e^{-5t}\sin(10\pi t)$, the two original signals and the added signal $y(t)$ is displayed on the figure 3(a). The sliding singular spectrum entropy of signal $y(t)$ is given on the figure 3(b). Obviously, we can give the same results as the above example. It is difficult to make out the impact signal from the original signal displayed on the figure 3(a), but we can obviously make out the impact signal from the processed signal $A(t)$ displayed on the figure 3(b).

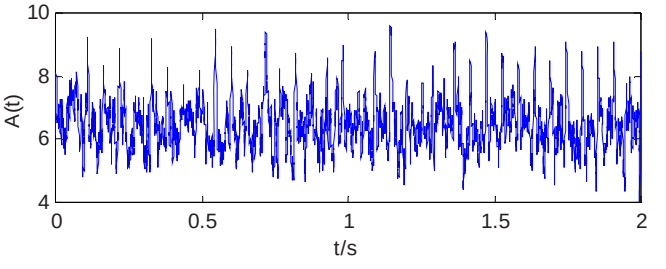
The weak impact signals mixed into the sine signal and chaos signal are acquired by the sliding singular spectrum entropy method in above examples. The abstracted weak impact signals are explicit by adopting the method, especially for the extraction of weak impact signal mixed in the chaos signal. The method is deduced based on the phase space reconstruction, in which, the time series is separated into different components by dividing the phase space into orthogonal sub-spaces. This can express the relative position of attractor in the higher dimensional space clearly, so it has good effects on the nonlinear and non-stationary signal.

4 Feature Extraction from Fault Gear Vibration Signals

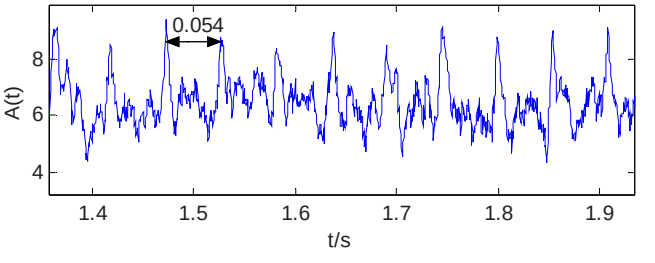
When the machinery parts such as gear and bearing are broken out, the vibration signals we gain from it are often mixed with periodic impact signals. From above simulation results, we know that the proposed method in this paper can abstract the hidden feature signals. So we use the method for the weak impact feature extraction on the machinery fault diagnosis. The experiment signals are collected from the fault experiment instrumentation for machinery driver system, which is made up of a motor, a reduction and a magnetic powder brake used to increase load. In the experiment, the driver system works on no-load. The reduction gear box is single stage for speed reduction. The drive gear wheel has 19 gears, and the follower gear has 95 gears. The speed of rotation for high speed shaft is 1100r/min. A gear in the drive gear has breakdown, the vibration signal generated by the gearbox was picked up by a CA-YD-103 acceleration sensors from the housing plate of the reduction gearbox, and it is acquired by a data acquisition system. The sampling frequency of the acceleration is 2000 Hz. The rotation frequency f_r of the drive gear shaft is 18.33Hz (f_r=1100/60=18.33Hz), from the equation, we can know that the interval of impact signal for the broken drive gear is 0.054 second. The original vibration waveform acquired by the acceleration sensors fitted on the housing plate is given on figure 4(a). Obviously, the signal mixes with strong background noise; the periodic impact signal



(a) The original vibration waveform of the reduction gear box



(b) The sliding singular spectrum entropy of the original vibration signal



(c) Local information of figure 4(b) from 1.4 second to 1.9 second

Fig. 4. The extraction of weak impact signal in gearbox vibration signal

is difficult to observe from it. The sliding singular spectrum entropy of the original vibration signal is given on figure 4(b); local information of figure 4(b) from 1.4 second to 1.9 second is given as figure 4(c). From which we can gain the periodic impact signal produced by the broken drive gear; the periodic is 0.054 second, which is consistent with the rotation cycle of the drive gear shaft.

5 Conclusions

A sliding singular spectrum entropy method based on the singular spectrum entropy is proposed, which is used for the singularity detection and the extraction of impaction signal. It also gives a method to find early symptoms of a potential failure in a gearbox and other components. The results show that the method can extract impact signal

hidden in the original time series and express the structure changing of time series effectively. From the results of above studies, we can conclude as followings.

1) The sliding singular spectrum entropy method can be used for the machinery online monitoring and fault diagnosis system. Because every point in time series is calculated with higher dimensional space reconstruction method and singular value decomposition, the parameter k can't be larger, which can lead to the calculation efficiency low, so the online calculation can't be carried out. Studies indicate that when the length of the original time series is 1024~2048, the parameter k is given value between 20 and 40, the feature extraction is valid.
2) The sliding singular spectrum entropy method combined with the traditional fault diagnosis method such as power spectrum method can identify the fault feature effectively, which gives a new path for the machinery fault diagnosis.
3) The method to calculate sliding singular spectrum entropy can be extended to the approximate entropy, sample entropy and wavelet entropy, then the sliding approximate entropy, sample entropy and wavelet entropy can be calculated, which gives a new method for the extraction of hidden feature in time series.

Acknowledgements

This work is supported by National Nature Science Foundation of China (No.50705069) and Science Foundation of Hubei education office (No.Q200611002).

References

1. Kanty, H., Schreiber, T.: Nonlinear Time Series Analysis. Cambridge University Press, Cambridge (1997)
2. Metin, A.: Nonlinear Biomedical Signal Processing. In: Dynamic Analysis and Modeling, vol. II. IEEE Press, New York (2001)
3. Yang, W.X., Jiang, J.S.: Study on the Singular Entropy of Mechanical Signal. Chinese Journal of Mechanical Engineering 36(12), 10–13 (2000)
4. Lu, Z.M., Zhang, W.J., Xu, J.W.: Application of Noise Reduction Method Based on Singular Spectrum in Gear Fault Diagnosis. Chinese Journal Of Mechanical Engineering 35(3), 95–99 (2000)
5. Yu, B., Li, Y.H., Zhang, P.: Application of Correlation Dimension and Kolmogorov Entropy in Aeroengine Fault Diagnosis. Journal of Aerospace Power 21(1), 22–24 (2006)
6. Shen, T., Huang, S.H., Han, S.M., Yang, S.Z.: Extracting Information Entropy Features for Rotating Machinery Vibration Signals. Chinese Journal of Mechanical Engineering 37(6), 94–98 (2001)
7. Roberts, S.J., Penny, W., Rezek, I.: Temporal and Spatial Complexity Measures for EEG-based Brain-computer Interfacing. Medical and Biological Engineering and Computing 37(1), 93–99 (1998)
8. Lu, Y., Li, Y.R., Wang, Z.G.: Research on a Extraction Method for Weak Fault Signal and Its application. Journal of Vibration Engineering 20(1), 24–28 (2007)
9. Liu, X.D., Yang, S.P., Shen, Y.J.: New Method of Detecting Abrupt Information Based on Singularity Value Decomposition and Its Application. Chinese Journal of Mechanical Engineering 38(6), 102–105 (2002)

10. Tenenbaum, J.B., Silva, V., Langford, J.C.: A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science 22, 2319–2323 (2000)
11. Schreiber, T., Schmitz, A.: On the Discrimination Power of Measures for Nonlinearity in a Time Series. Phys. Rev. E. 55, 5443–5447 (1997)
12. Kim, K., Parlos, A.G.: Model-based Fault Diagnosis of Induction Motors Using Non-stationary Signal Segmentation. Mechanical Systems and Signal Processing 16(2-3), 223–253 (2002)
13. Cheng, J., Yu, D., Yang, Y.: Application of an Impulse Response Wavelet to Fault Diagnosis of Roller Bearings. Mechanical Systems and Signal Processing 21(2), 920–929 (2007)
14. Abbasion, S., Rafsanjani, A., Farshidianfar, A., et al.: Rolling Element Bearings Multi-fault Classification Based on the Wavelet De-noising and Support Vector Machine. Mechanical Systems and Signal Processing 21(7), 2933–2945 (2007)
15. Lu, Y., Li, Y.R., Xu, J.W.: Weighted Phase Space Reconstruction Algorithm for Noise Reduction and its Application. Chinese Journal of Mechanical Engineering 43(8), 171–174 (2007)

Security Assessment Framework Using Static Analysis and Fault Injection

Hyungwoo Kang

Financial Supervisory Service
27 Yoido-dong, Youngdeungpo-gu, Seoul, 150-743, Korea
kanghw@fss.or.kr

Abstract. For large scale and residual software like network service, reliability is a critical requirement. Recent research has shown that most of network software still contains a number of bugs. Methods for automated detection of bugs in software can be classified into static analysis based on formal verification and runtime checking based on fault injection. In this paper, a framework for checking software security vulnerability is proposed. The framework is based on automated bug detection technologies, i.e. static analysis and fault injection, which are complementary each other. The proposed framework provides a new direction, in which various kinds of software can be checked its vulnerability by making use of static analysis and fault injection technology. In experiment on proposed framework, we find unknown vulnerability as well as known vulnerability in Windows network module.

Keywords: Security assessment, Static analysis, Fault injection, RPC (Remote Procedure Call), Software security, Buffer overflow.

1 Introduction

A real-world system is always likely to contain faults. There are many efforts to try to remove the faults, but unknown vulnerabilities are reported without a break. For large scale and residual software like network service, reliability is a critical requirement. It is thus desirable to have automated bug detection methods which utilize the vast power of machines available today to reduce human testing time. Methods for automated detection of bugs in software can be classified into static analysis based on formal verification and dynamic analysis based on fault injection. In this paper, a framework for checking software security vulnerability is proposed. The framework is based on, automated bug detection technologies, static analysis and fault injection, which are complementary each other. There are lots of previous works related to static analysis and fault injection. However, those works are partial researches about special type of software in order to detect security vulnerability. The proposed framework can check security vulnerability for most types of software.

This paper is organized as follows. Chapter 2 reviews previous research relating static analysis and fault injection. In chapter 3, we propose a framework for checking software security vulnerability. Experimental result is presented in chapter 4. Finally, in chapter 5, we draw conclusions.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 679–687, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2 Related Works

2.1 Static Analysis

Previous static analysis techniques can be classified into the following two types: lexical analysis based approach and semantic based approach.

2.1.1 Lexical Analysis Based Approach

Lexical analysis technique is used to turn the source codes into a stream of tokens, discarding white space. The tokens are matched against known vulnerability patterns in the database. Well-known tools based on lexical analysis are Flawfinder [1], RATS [2], and ITS4 [3]. For example, these tools find the use of vulnerable function like *strcpy*() by scanning source codes. While these tools are certainly a step up from UNIX utility *grep*, they produce a hefty number of false positives because they make no effort to account for the target code's semantics.

To overcome the weakness of lexical analysis approach, a static analysis tool must leverage more compiler technology. Abstract syntax tree (AST) analysis approach [4] finds security vulnerability pattern by building an AST from source code.

2.1.2 Semantic Based Approach

BOON [5] applies integer range analysis to detect buffer overflows. The range analysis is flow-insensitive and generates a very large number of false alarms. CQUAL [6] is a type-based analysis tool that provides a mechanism for checking properties of C programs. It has been used to detect format string vulnerability. SPLINT [7] is a tool for checking C program for vulnerabilities and programming mistakes. It uses a lightweight data-flow analysis to verify assertions about the program.

An abstract interpretation approach [8, 9] is used for verifying the AirBus software. Although this technique is a mature and sound mathematical approach, the static analysis tool that uses abstract interpretation can't scale to larger code segments. According to the testing paper [10], target program has to be manually broken up into 20-40k lines-of-code blocks to use the tool.

Microsoft set up the SLAM [11, 12, 13] project that uses software model checking to verify temporal safety properties in programs. It validates a program against a well designed interface. However, SLAM does not yet scale to very large programs because of considering data flow analysis. MOPS [14, 15] is a tool for verifying the conformance of C programs to rule that can be expressed as temporal safety properties. It represents these properties as finite state automata and uses a model checking approach to find if any insecure state is reachable in the program.

2.2 Fault Injection

Fault injection has been proposed to be used to understand the effects of real faults, to get feedback for system correction or enhancement and to forecast the expected system behavior. A variety of techniques have been proposed. Among them, software-implemented fault injection (SWIFI) [16] has recently risen to prominence. MAFALDA [17] uses an experimental approach based on SWIFI targeting both external and internal faults for microkernel-based systems. Fuzz [18] and Ballista [19, 20] are research project that have

studied the robustness of UNIX system software. Ballista performs fault injection at the API level by passing combinations of acceptable and exceptional inputs as a parameter list to the module under test via an ordinary function call, while Fuzz feeds completely random input to applications without using any model of program behavior. The Wisconsin team extends their research to the robustness of Windows NT application [21] in 2000. These researches and practices prove the effectiveness of fault oriented software robustness assessment. In a paper [22], published in 2002, Dave Aitel describes an effective method, called Block-Based Protocol Analysis, implemented in SPIKE [23]. Protocols can be decomposed into length fields and data fields. It can reduce the size of the potential space of inputs. Kang and Lee [24] provide a security assessment methodology for improving reliability of network software with no published protocol specification. The methodology improves the efficiency of security assessment for application network services by feeding fault data to the parameter field of a target function in network packet. A good practice has been conducted by PROTOS project [25], which adopted an interface fault injection approach for security testing of protocol implementations. Surprisingly, it successfully reported various software flaws in a few protocol implementations including SNMP, HTTP and SIP, many of which have been in production for years. There are software testing tools using fault injection such as Holodeck [26, 27] and Mangleme [28]. Holodeck focuses on network faults, disk faults and memory faults. The tool finds faults that general program occur, so it is hard to perform security testing that focuses on the specific application such as web browser. The Mangleme finds the vulnerability of browser by inserting various data.

2.3 Comparison of Static Analysis and Fault Injection

2.3.1 Advantages of Static Analysis

First, while fault injection tools are only useable when a running system must be given to them, static analysis does not require it. If users with the static analysis tools have the source code of the software, they are willing to examine the robustness of software without running it. But fault injection can only check a system available to the user.

Second, although static analysis may need to filter through falsely positive error results from static analysis tests, static analysis can point out errors in software much more precisely than fault injection.

Third, the tests done by fault injection can never be as exhaustive and comprehensive as those done by static analysis. It is simply understandable fact of the nondeterministic nature of program execution that paths of certain code may never be utilized by any testing tool even though this is the main goal of fault injection tools. As a result some errors may go undetected because of this reason. Static analysis does not suffer from this matter because it always checks all of the code made available to it.

2.3.2 Advantages of Fault Injection

First, fault injection does not require source code to be present, while static analysis clearly does. Although this capability does not support a user to improve the software, it does allow evaluation and verification of functionality for a particular purpose. Fault injection is also independent of source language because it does not depend on source.

Second, there is a difference of scale between static analysis and fault injection in relation to program size. It is possible for static analysis to analyze the software within the scope of the size of the source. In the case of fault injection, the performance is independent of source size, but it is dependent on how long the program takes to run or how many fault injection scenarios the tester decides to run on it.

A third advantage of fault injection relates to its stated purpose of verifying error recovery methods in software package. Static analysis may be able to determine whether or not a program tests for and attempts to handle a particular error, but it has no way of determining if the error handling code will be successful. As a result, fault injection can test that software recover a program to life by correctly coping with the error, but static analysis cannot.

3 Security Assessment Framework Using Static Analysis and Fault Injection

3.1 Framework

Static analysis and fault injection are willing to be combined into one framework because they have advantages being able to make up for their weak points each other. The drawback of static analysis is that the number of false positive can be very large and it can only test the vulnerability of software with a source code. Although static analysis is useful to do an invariant checking, a formal verification, and a code style analysis, it is obviously also impractical if a large program is target software of the vulnerability analysis. Fault injection has an ability to examine an automated black box testing and verify the recovery method verification. But the vulnerability analysis with fault injection may result in large numbers of false negative.

Figure 1 shows the proposed framework for software security assessment by making use of static analysis technology and fault injection technology. The process of security assessment is shown as follow:

- Step 1 : Selection of target software
- Step 2 : Decision of security assessment technology by making use of checklist
- Step 3 : Validation process of target software by making use of static analysis
  - Input the source code of target software and security property
  - Execution of static analysis tool for checking security vulnerability
  - Analysis of alarm produced from static analysis tool
- Step 4 : Validation process of target software by making use of fault injection
  - Execution of target binary code
  - Injection of fault data, which consist of random data and semivalid data, to the interface of target software being executed
  - Analysis of response to fault injection
- Step 5 : Removal of the vulnerability of target software

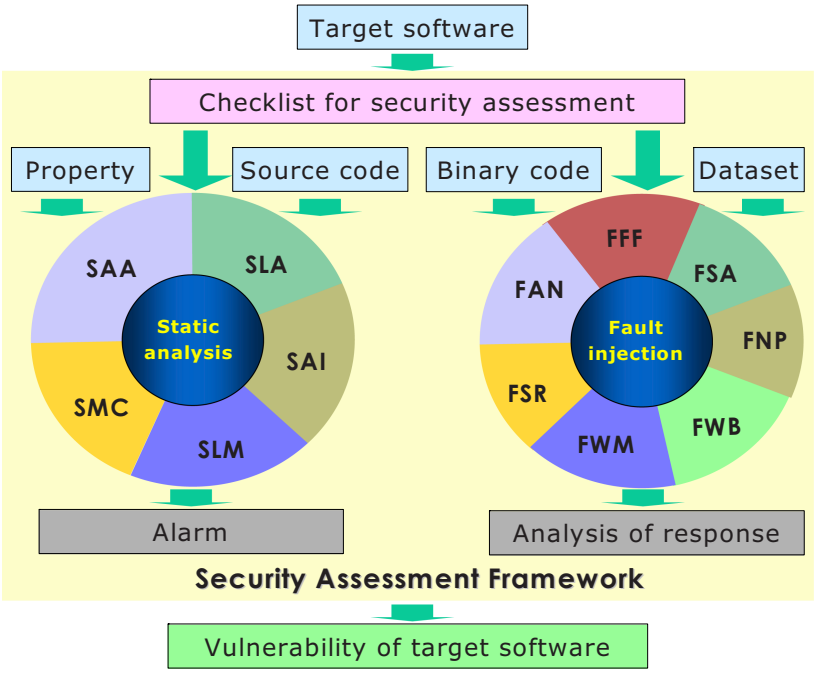


Fig. 1. Security assessment framework using static analysis and fault injection

There are lots of previous works related static analysis and fault injection. However, these works are partial researches about special type of software in order to detect software security vulnerability. In other words, there is no research about several type of software. In this paper, we present a framework for software security assessment, which can cover most type of software by complement of several security assessment method.

Table 1. Meaning of abbreviated word used in figure 1

Technology	Abbreviated word	Meaning
Static analysis	SLA	Lexical Analysis technology
	SAA	AST Analysis technology
	SLM	Light-weight Model checking technology
	SMC	Model Checking technology
	SAI	Abstract Interpretation technology
Fault injection	FAN	Application Network service fault injection
	FSR	System Resource fault injection
	FWM	Windows Message & event fault injection
	FWB	Web Browser fault injection
	FNP	Network Protocol fault injection
	FSA	System API fault injection
	FFF	File Format fault injection

The meaning of abbreviated word used in figure 1 is shown at table 1. The first character 'S' and 'F" of the abbreviated words means the static analysis technology and fault injection technology, respectively.

3.2 Static Analysis Based Method

Static analysis is an important assessment technology since the use of static analysis tools promises to provide the last major hurdles in prevention of logical bugs and code problems in software development. Static analysis technology contributes to analyze not only the syntax, but also the semantics of a program. While it is possible to build the more complex analysis of a program by utilizing static analysis, this technology still needs user interaction due to its false positives. Static analysis based methods are classified into following five kinds of methods according to size of target software and type of vulnerability which we try to find in target software.

SLA checks whether there is a usage of vulnerable function in target software. SAA checks whether there is a usage of vulnerable pattern in target software. SLM checks temporal safety property without data flow analysis. SMC checks temporal safety property with data flow analysis. SAI abstracts the semantic of program and checks boundary of buffer used in target software. The more we use low-layer method, the more we can analyze the target software in detail. And, the more we use high-layer method, the more we can analyze large scale software.

Table 2 shows major researches relating static analysis based methods. We can use the major researches in case of security assessment.

Table 2. Major researches relating static analysis based methods

Method	Description	Major research
SLA	Lexical analysis	Flawfinder [1], RATS [2], ITS4 [3]
SAA	AST analysis	AST analysis[4]
SLM	Light-weight model checking	MOPS [14, 15]
SMC	Model checking	SLAM [11, 12, 13]
SAI	Abstract interpretation	Polyspace [8, 9]

3.3 Fault Injection Based Method

Fault injection is also a useful technology in developing high quality and reliable code. Its ability to reveal how software systems behave under experimentally controlled anomalous circumstances makes it an ideal crystal ball for predicting how badly good software can behave. There are a number of different methods of fault injection, but all methods attempt to cause the execution of seldom used control pathways within a system. However, fault injection technology inherently has the drawback that the test with them may result in large numbers of false negatives.

Fault injection based methods are classified into following 7 kinds of methods according to the type of target software. Fault injection methods are classified into 5 types according to interface types of software. Once target software is chosen to validate security vulnerability, we can use more than one types of target software according to its interface type.

FAF is suitable for security assessment for application program interface(API) exported by target software. FNP can check security vulnerability of network protocol publishing its specification. FAN is for security assessment for improving reliability of network software without publishing protocol specification. FWB is for checking security vulnerability of web browser like Internet Explorer. FWM is suitable for Windows application program having lots of message transmission. FSR is for security assessment of software using system resource such as, memory and disk. FFF is suitable for checking security vulnerability of software having interface of file type.

Table 3 shows major researches relating fault injection based methods. We can use the major researches in case of security assessment.

Table 3. Major researches relating fault injection based methods

Method	Type of target software	Major research
FAN	Application network service	SPIKE [23], Kang and Lee[24]
FSR	System resource	Holodeck [26, 27]
FWM	Windows message & event	Message fuzzer [21]
FWB	Web browser	Mangleme [28]
FNP	Network protocol	PROTOS project [25]
FSA	System API	Fuzz [18], Balista [19, 20]
FFF	File format	File Fuzzer [29]

4 Experiment

We make experiment on Windows RPC network services, which are RPCSS service and browser service, as the target software for security assessment. These target programs are not small size and use network protocol unpublished its specification. Therefore, we make security assessment for target network service using SAA method (AST analysis technology) [4] and FAN method (application network service fault injection technology) [24] in proposed framework.

We find a known vulnerability [30] in RPCSS service module by making use of SAA method. The vulnerability can cause a stack buffer overflow in target system. And moreover, We find an unknown vulnerability causing DoS in SMB module by making use of FAN method when we make a experiment for "BrowserrSetNetlogonState()" function in browser service. A target system exposed under the vulnerability does not accept SMB connection using 445 port from remote system. No fault is found in target system during experiment on the rest of the functions.

5 Conclusion

We proposed a security assessment framework for software. We can analyze the vulnerability of software more efficiently and correctly that using the framework combined with static analysis and fault injection. The reason why it is possible that a tool combining the two approaches could prove valuable to software developers because they have the same determined goal. The strength of the two types of tools can be

combined in mutually advantageous ways. The assessment framework combined with two technologies give not only the arguably lower user interaction cost of static analysis but also the lower probability that fault injection might not even detect certain programming errors.

An experiment on proposed framework was provided. Windows RPC network services were chosen as target software for the experiment. As we mentioned in the chapter 4, we found known vulnerability causing stack buffer overflow in RPCSS module and unknown vulnerability causing DoS in browser service module.

Our immediate research results include the setting-up of framework being able to validate the reliability of software. Finally natural extensions to this work will be to add results of experiment on various kinds of software.

References

1. Wheeler, D.A.: Flawfinder, http://www.dwheeler.com/flawfinder/
2. RATS, http://www.securesw.com/rats/
3. Viega, J., Bloch, J.T., Kohno, T., McGraw, G.: ITS4: A Static Vulnerability Scanner for C and C++ Code. ACM Transactions on Information and System Security 5(2) (2002)
4. Kang, H., Kim, K., Hong, S., Lee: A Model for Security Vulnerability Pattern. In: Gavrilova, M.L., Gervasi, O., Kumar, V., Tan, C.J.K., Taniar, D., Laganá, A., Mun, Y., Choo, H. (eds.) ICCSA 2006. LNCS, vol. 3982, pp. 385–394. Springer, Heidelberg (2006)
5. Wagner, D., Foster, J.S., Brewer, E.A., Aiken, A.: A First Step towards Automated Detection of Buffer Overrun Vulnerabilities. In: Network and distributed system security symposium, San Diego, CA, pp. 3–17 (2000)
6. Foster, J.: Type qualifiers: Lightweight Specifications to Improve Soft-ware Quality. Ph.D. thesis, University of California, Berkeley (2002)
7. Evans, D.: SPLINT, http://www.splint.org/
8. Blanchet, B., Cousot, P., Cousot, R., Feret, J., Mauborgne, L., Mine, A., Monniaux, D., Rival, X.: A Static Analyzer for Large Safety-Critical Software (2003)
9. Abstract interpretation (2001), http://www.polyspace.com/downloads.htm
10. Zitser, M., Lippmann, R., Leek, T.: Testing Static Analysis Tools using Exploitable Buffer Overflows from Open Source Code. In: SIGSOFT 2004, pp. 97–106 (2004)
11. Ball, T., Majumdar, R., Millstein, T., Rajamani, S.: Automatic Predicate Abstraction of C Programs. PLDI. ACM SIGPLAN Not. 36(5), 203–213 (2001)
12. Ball, T., Podelski, A., Rajamani, S.: Relative Completeness of Abstraction Refinement for Software Model Checking. In: Katoen, J.-P., Stevens, P. (eds.) TACAS 2002. LNCS, vol. 2280, pp. 158–172. Springer, Heidelberg (2002)
13. Ball, T., Rajamani, S.: The SLAM project: Debugging System Software via Static Analysis. In: 29th ACM POPL. LNCS, vol. 1254, pp. 72–83. Springer, Heidelberg (2002)
14. Chen, H., Wagner, D.: MOPS: An Infrastructure for Examining Security Properties of Software. In: Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS), Washington, DC (2002)
15. Chen, H., Wagner, D., Dean, D.: Setuid Semystified. In: Proceedings of the Eleventh Usenix Security Symposium, San Francisco, CA (2002)
16. Voas, J.M., McGraw, G.: Software Fault Inoculating Programs Against Errors. Wiley Computer Publishing, Chichester
17. Fabre, J.C., Rodriguez, M., Arlat, J., Sizun, J.M.: Building Dependable COTS Microkernel-based Systems using MAFALDA. In: Pacific Rim International Symposium on Dependable Computing (PRDC 2000), pp. 85–92 (2000)

18. Miller, B.P., Fredriksen, L., So, B.: An Empirical Study of the Reliability of UNIX Utilities. Communications of the ACM 33(12) (1990)
19. Koopman, P., Sung, J., Dingman, C., Siewiorek, D., Marz, T.: Comparing Operating Systems using Robustness Benchmarks. In: 16th IEEE Symposium on Reliable Distributed Systems, pp. 72–79 (1997)
20. Kropp, N.P., Koopman, P.J., Siewiorek, D.P.: Automated Robustness Testing of Off-the-Shelf Software Components. In: 28th International Symposium on Fault- Tolerant Computing, pp. 464–468 (1998)
21. Justin, E.F., Barton, P.M.: An Empirical Study of the Robustness of Windows NT Applications using Random Testing, http://www.cs.wisc.edu/_bart/fuzz/fuzz.html
22. Aitel, D.: The Advantages of Block-based Protocol Analysis for Security Testing (2002), http://www.immunitysec.com/resources-papers.shtml
23. SPIKE Development Homepage, http://www.immunitysec/spike.html
24. Kang, H., Lee, D.: Security Assessment for Application network Service using Fault Injection. In: Yang, C.C., Zeng, D., Chau, M., Chang, K., Yang, Q., Cheng, X., Wang, J., Wang, F.-Y., Chen, H. (eds.) PAISI 2007. LNCS, vol. 4430, pp. 172–183. Springer, Heidelberg (2007)
25. PROTOS: Security Testing of Protocol Implementation, http://www.ee.oulu.fi/research/ouspg/protos
26. Holodeck, http://www.securityinnovation.com/
27. James, A.W., Herbert, H.T.: How to Break Software Security. Addison Wesley, Reading
28. Mangleme, http://freshmeat.net/projects/managleme
29. Michael, S., Adam, G.: The Art of File Format Fuzzing, Blackhat, USA (2005)
30. Microsoft Security Bulletin MS03-026. Microsoft (2003), http://www.microsoft.com/technet/security/bulletin/MS03-026.mspx

Improved Kernel Principal Component Analysis and Its Application for Fault Detection

Chuyao Chen, Daqi Zhu, and Qian Liu

Information Engineering College, Shanghai Maritime University,
200135 Shanghai, China
zdq367@yahoo.com.cn

Abstract. The kernel principal component analysis (KPCA) based on feature vector selection (FVS) is proposed in this paper for fault detection in nonlinear system. Firstly, the KPCA algorithm is described in detail. Secondly, a feature vector selection (FVS) scheme based on a geometric consideration is adopted to reduce the computational cost of KPCA. Finally, the KPCA and KPCA based on FVS (FVS-KPCA) are applied to a simple nonlinear system. The fault detection results and the comparison confirm the superiority of FVS-KPCA in fault detection.

Keywords: kernel principal component analysis (KPCA); feature vector selection (FVS); fault detection.

1 Introduction

Numerous PCA-based statistical process monitoring methods combining with T^2 charts, Q charts and contribution plots [1] [2]have been proposed and applied to fault detection, but the applications of PCA-based monitoring to inherently nonlinear processes may lead to inefficient and unreliable results. Because a linear PCA is inappropriate to describe the nonlinearity within the process variables[3].To cope with this problem, extended versions of PCA suitable for handling nonlinear systems have been developed, such as methods based on neural network [4][5][6], methods based on principal curve [7], methods based on kernel function [8] [9].In recent years, KPCA has been utilized directly for nonlinear process monitoring. Unlike nonlinear methods based on neural networks, KPCA does not involve a nonlinear optimization procedure, and the calculation in linear PCA can be directly used in KPCA. But one drawback of KPCA is that the computation time may increase with the number of samples. To solve this problem, we propose the algorithm of KPCA based on FVS, and highlight the power of kernel PCA as a nonlinear process monitoring technique.

2 KPCA

Conceptually, KPCA consists of two steps: (1) the mapping of the data set from its original space into a hyper-dimensional feature space F via a nonlinear

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 688–695, 2008.
© Springer-Verlag Berlin Heidelberg 2008

mapping ϕ, and (2) a PCA calculation in F. In implementation, the implicit feature vector in F does not need to be computed explicitly, while it is just done by computing the inner product of two vectors in F with a kernel function.

Let $x_1, x_2, \ldots, x_n \in R^m$ be the n training samples (column vectors) for KPCA learning. By the nonlinear mapping ϕ, the measured inputs are extended into the hyper-dimensional feature space F as follows

$$\phi : x \in R^m \rightarrow \phi(x) \in F^h \ . \tag{1}$$

The dimension of the feature space,h, can be arbitrary large or even infinite. The mapping of x_i is noted as $\phi(x_i) = \phi_i$. Where column vector $m_\phi = \sum_{i=1}^{n} \phi_i / n$ is the sample mean. The sample covariance in F can be constructed by

$$C_\phi = \frac{1}{n} \sum_{i=1}^{n} (\phi_i - m_\phi)(\phi_i - m_\phi)^T \ . \tag{2}$$

It is easy to see that every eigenvector ν of C_ϕ can be linearly expanded by

$$\nu = \sum_{i=1}^{n} \alpha_i \phi_i \ . \tag{3}$$

To obtain the coefficients $\alpha_i (i = 1, 2, \ldots, n)$, a kernel matrix K with size $n \times n$ is defined, whose elements are determined by virtue of kernel tricks

$$K_{ij} = \phi_i^T \phi_j = (\phi_i \cdot \phi_j) = k(x_i, x_j) \ . \tag{4}$$

Here,$k(x_i, x_j)$ is the calculation of the inner product of two vectors in F with a kernel function. Centralize K by

$$\overline{K} = K - 1_n K - K 1_n + 1_n K 1_n \ . \tag{5}$$

Here, matrix $1_n = (1/n)_{n \times n}$.

Let column vectors $\gamma_i (i = 1, 2, \ldots, p; 0 < p \leq n)$ be the orthonormal eigenvectors of $\overline{K}$ corresponding to the largest p positive eigenvalues,$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$, and p is the number of principal components in KPCA. The orthonormal eigenvectors $\beta_i (i = 1, 2, \ldots, p)$ of $C_\phi$corresponding to the largest p positive eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_p$, are

$$\beta_i = \frac{1}{\sqrt{\lambda_i}} (\phi_1, \phi_2, \ldots, \phi_n) \gamma_i \ . \tag{6}$$

To a new column-vector sample x_{new}, the mapping to the feature space is $\phi(x_{new})$. The projection of x_{new} onto the eigenvectors $\beta_i (i = 1, 2, \ldots, p)$ is $t = (t_1, t_2, \ldots, t_p)^T$, and it is also called KPCA-transformed feature vector

$$t = (\beta_1, \beta_2, \ldots, \beta_p)^T \phi(x_{new}) \ . \tag{7}$$

The ith KPCA-transformed feature t_i can be obtained by

$$\begin{aligned} t_i &= \beta_i^T \phi(x_{new}) \\ &= \frac{1}{\sqrt{\lambda_i}} \gamma_i^T (\phi_1, \phi_2, \ldots, \phi_n)^T \phi(x_{new}) \\ &= \frac{1}{\sqrt{\lambda_i}} \gamma_i^T [k(x_1, x_{new}), k(x_2, x_{new}), \ldots, k(x_n, x_{new})]^T, i = 1, 2, \ldots, p \ . \end{aligned} \tag{8}$$

By using Eq. (8), the KPCA-transformed feature vector of a new sample vector $t_{new} = [t_1, t_2, \ldots, t_p]$ can be obtained.

3 FVS-KPCA

However, KPCA has its shortcomings to limit its application. In the training stage of KPCA, it requires to store and manipulate the kernel matrix K, the size of which is the square of the number of samples. When the sample number becomes large, the calculation of eigenvalues and eigenvectors will be time consuming. To solve this problem, a feature selection scheme is adopted to reduce the computational complexity of KPCA. The concrete algorithm is fully described in the literature [10].

In Eq.(3), all the training samples in F, $\phi_i (i = 1, 2, \ldots, n)$, are used to present eigenvector ν. In practice, the dimensionality of the subspace spanned by ϕ_i is just equal to the rank of kernel matrix K, and the rank of K is often less than n, that is $rank(K) < n$. Especially, when the training data set is very large, $rank(K) \leq n$ [11] [12].

Suppose a basis of the feature vectors,$\phi_{b_i} (i = 1, 2, \ldots, rank(K); b_i \in [1, n])$, is known, then Eq.(3) can be written as follows

$$\nu = \sum_{i=1}^{rank(K)} \alpha_{b_i} \phi_{b_i} \ . \tag{9}$$

Since $rank(K) \leq n$, this will largely reduce the computational complexity.

In order to obtain such a basis of the feature vectors in F, a geometric consideration based method [11] is used. The idea is to look for a subset of the samples whose mappings in F are sufficient to express all the data in F as a linear combination of them.

Assume that $S = x_{S_1}, x_{S_2}, \ldots, x_{s_{L_S}}$ is a selected sample set, where L_S is the number of the selected vectors. The estimation of the mapping of any vector x_i is regarded as a linear combination of S in F, and it can be written as follows

$$\widehat{\phi_i} = \varnothing_S \cdot \tau_i \ . \tag{10}$$

where $\varnothing_S = (\varnothing_{S_1}, \varnothing_{S_2}, \ldots, \varnothing_{S_{L_S}})$ is the mapping matrix of the selected samples in F and $\tau_i = (\tau_i^1, \tau_i^2, \ldots, \tau_i^{L_S})$ is the coefficient vector. Then the question becomes how to find the coefficient vector τ_i , so as to make the estimated mapping $\widehat{\phi_i}$ as close to the real mapping ϕ_i as possible. This can be achieved by minimizing the following equation

$$\delta_i = \frac{\left\| \phi_i - \widehat{\phi_i} \right\|^2}{\left\| \phi_i \right\|^2} \ . \tag{11}$$

Putting the derivatives with respect to τ_i to zero, i.e. $\frac{\partial \delta_i}{\partial \tau_i} = 0$, and rewriting it in matrix form, yields

$$\min(\delta_i) = 1 - \frac{K_{Si}^T K_{SS}^{-1} K_{Si}}{k_{ii}} \ . \tag{12}$$

Where $k_{ii} = k(x_i, x_i)$,$K_{SS} = (k(x_{S_p}, x_{S_q}))_{1 \le p \le L_S, 1 \le q \le L_S}$ is a square matrix of dot product of the selected vectors, and $K_{Si} = (k(x_{S_p}, x_{S_i}))_{1 \le p \le L_S}$ is the vector of dot product between x_i and the selected set S.

Letting Eq.(12) satisfy all the samples, we can get

$$\min_S J_S = \frac{1}{n} \sum_{x_i} \frac{K_{Si}^T K_{SS}^{-1} K_{Si}}{k_{ii}} \ . \tag{13}$$

Solution of the problem can be obtained by an iterative process, and the process stops when K_{SS} is no longer invertible or a predefined number of selected vectors is reached [11].

That is how we get the selected sample vector set $S = x_{S_1}, x_{S_2}, \ldots, x_{s_{L-S}}$. When S is adopted in the KPCA algorithm, the computational complexity is reduced. When a new sample x_{new} is available, the ith KPCA-transformed feature vector can be calculated using Eq.(9) as follow

$$t_i = \frac{1}{\sqrt{\lambda_i}} \gamma_i^T [k(x_{S_1}, x_{new}), k(x_{S_2}, x_{new}), \ldots, k(x_{S_{L_S}}, x_{new})]^T \ . \tag{14}$$

4 Fault Detection Based on KPCA and FVS-KPCA

Similar to T^2 statistics and squared prediction error (SPE) statistics in linear PCA, the corresponding metrics in the feature space of KPCA and FVS-KPCA are defined as follows

$$T_f^2 = [t_1, t_2, \ldots, t_p] \Lambda^{-1} [t_1, t_2, \ldots, t_p]^T \ . \tag{15}$$

where $Lambda^{-1}$ is the diagonal matrix of the inverse of the eigenvalues associated with the retained principal components.$SPE_f$is calculated as it is in literature [8]

$$SPE_f = \overline{K}(x_{new}, x_{new}) - t_{new}^T t_{new} \ . \tag{16}$$

T_f^2 and SPE_f are used for nonlinear fault detection, that is, when a fault occurred, the T_f^2 and SPE_f changed abruptly.

5 Simulations and Comparisons

In order to validate the fault detection performance of KPCA, the system used by literature [3] is simulated. It consists of three variables nonlinearly related with t, which are generated by

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ t^2 - 3 \\ -t^2 + 3t^2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \ . \tag{17}$$

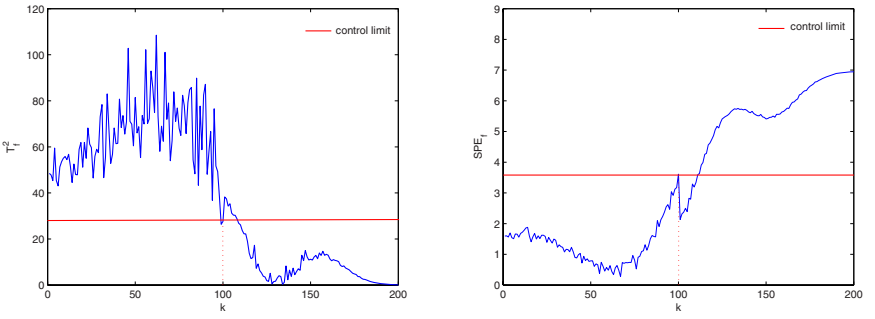


Fig. 1. T_f^2 and SPE_f for fault detection based on KPCA

where $e_i \sim N(0, 0.01)$ is a white Gaussian noise and $t \in [0.01, 1]$. The normal training data comprise 100 samples (samples $1 \sim 100$) obtained using Eq. (17). The test data set consists of two parts: the first 100 samples (samples $101 \sim 200$) are collected in the normal condition in training data; and the second 100 testing samples (samples $201 \sim 300$) are obtained using the system but with x_3 expressed as $x_3 = -1.1t^3 + 3.2t^2 + e_3$ .Gaussian kernel function is used

$$k(x, y) = exp(-\frac{\|x - y\|^2}{\sigma}) \,. \tag{18}$$

where $\|\cdot\|$ is l_2-norm and σ is the width of a Gaussian function. The choice of σ significantly impacts the process monitoring performance. For detecting the very small deviations in nonlinear systems, we should set the width of the Gaussian function to be small, which will cause the T_f^2 statistic to have a relatively low value during the faulty condition compared to its average normal value [6].

First, We use KPCA to carry out fault detection of the nonlinear system above using $\sigma = 0.047$.Fig.1 shows the fluctuation of T_f^2 and SPE_f in the monitoring process. When a fault occurred ($k = 101$), T_f^2 falls below the control limit because the very small deviations in the nonlinear system,whereas SPE_f increases abruptly.

Then, FVS-KPCA is adopted to the nonlinear system expressed by Eq.(17) for fault detection. Here,$\sigma = 0.08$, $L_S = 20$. Fig.2 shows the fluctuation of T_f^2 and SPE_f based on FVS-KPCA, from which we can tell the difference of effectiveness and efficiency between KPCA and FVS-KPCA.

Compared to the results of KPCA, firstly, FVS-KPCA has nearly the same fault detection results as KPCA methods, and it is much more accurate. That's because the training stage in FVS-KPCA is completed in a short time, the nonlinear relations between the variables can be catched quickly and applied to fault detection immediately, which cause less delay in online monitoring system. As a result, the T_f^2 and SPE_f abruptly changes at the onset of the fault ($k = 101$).

Secondly, the computational cost of training process is reduced by FVS-KPCA, so that the efficiency of fault detection is increased compared to the traditional KPCA, especially when the training sample set is very large. Table.1

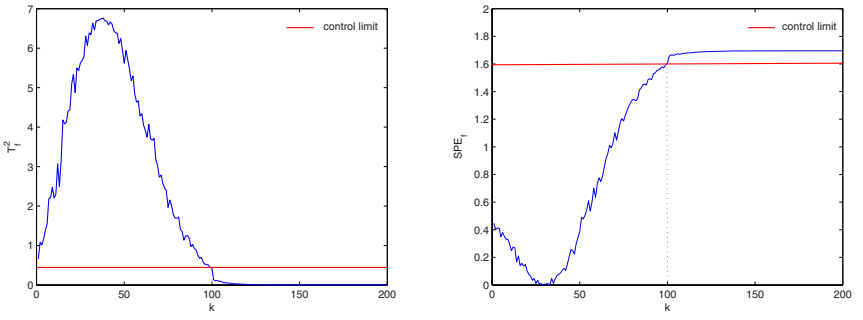


Fig. 2. T_f^2 and SPE_f for fault detection based on FVS-KPCA

Table 1. The time consumed by KPCA and FVS-KPCA in training stage

Num. of Samples	100	200	300	400	500
KPCA	0.2557s	0.3898s	0.7515s	1.2405s	2.1059s
FVS-KPCA	0.2103s	0.3461s	0.3945s	0.5805s	0.8387s
Num. of Samples	600	700	800	900	1000
KPCA	3.1732s	4.6135s	6.3530s	8.6048s	11.3345s
FVS-KPCA	1.1124s	1.9785s	2.5720s	3.2588s	4.5889s

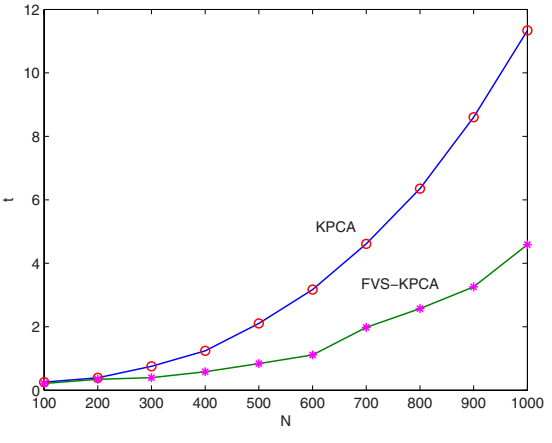


Fig. 3. The time comsumed by KPCA and FVS-KPCA

shows the time taken respectively by KPCA and FVS- KPCA in process monitoring of the same simulation sample described above.

Fig.3 gives us an obvious description of the time consumed by KPCA and FVS-KPCA. It's not hard to find that when the number of samples is relatively small, $n = 100$ or $n = 200$, for example, the time taken by KPCA and

FVS-KPCA in process monitoring is nearly the same. But as the number of samples increases, for example, $n = 900$ or $n = 1000$, KPCA is much more time-consuming than FVS-KPCA. The difference derives from the training process of KPCA and FVS-KPCA.

In nonlinear process monitoring, in order to keep the accuracy of fault detection, the training samples increase with the total sample number. In the algorithm of KPCA, all the training samples are used to obtain the kernel matrix so that the calculating process is time-wasting. Whereas in FVS-KPCA, a subset of the samples is found to instead of all training samples, the number of the selected vectors is less than the number of training vectors, especially when the training set is very large, so it's time-saving and efficient. That's why FVS-KPCA is much more attractive in process monitoring than KPCA.

6 Conclusions

This paper focuses on the improvement of KPCA algorithm and its application for fault detection of nonlinear multivariate process. Simulation results show that FVS-KPCA is superior than KPCA in fault detection, because the computation complexity is reduced significantly especially when the number of samples is very large.

Acknowledgments. This project is supported by the Natural Science Foundation of Shanghai (NO.07ZR14045), the Project of Shanghai Municipal Education Commission (NO.06FZ038) and National Key Technologies R&D Program(2007BAF10B05).

References

1. Kresta, J., MacGregor, J.F., Marlin, T.E.: Multivariate statistical monitoring of process operating performance. Canadian Journal of Chemical Engineering 69, 35–47 (1991)
2. Wise, B.M., Ricker, N.L.: Recent advances in multivariate statistical process control: improving robustness and sensitivity. In: Proceedings of IFAC International Symposium. ADCHEM 1991, pp. 125–130. Paul Sabatier University, Toulouse (1991)
3. Dong, D., McAvoy, T.J.: Nonlinear principal component analysis-based on principal curves and neural networks. Computers and Chemical Engineering 20, 65–78 (1996)
4. Kramer, M.A.: Non-linear principal component analysis using autoassociative neural networks. AIChE Journal 37, 233–243 (1991)
5. Tan, S., Mavrovouniotis, M.L.: Reducing data dimensionality through optimising neural networks inputs. AIChE Journal 41, 1471–1480 (1995)
6. Jia, F., Martin, E.B., Morris, A.J.: Non-linear principal component analysis for process fault detection. Computers and Chemical Engineering 22, S851–S854 (1998)
7. Hastie, T.J., Stuetzle, W.: Principal curves. Journal of American Statistical Association 84, 502–516 (1989)

8. Choi, S.W., Lee, J.M., Park, J.H., et al.: Fault detection and identification of nonlinear processes based on kernel PCA. Chemometrics and intelligent laboratory systems 75, 55–67 (2005)
9. Choi, S.W., Lee, I.B.: Nonlinear dynamic process monitoring based on dynamic KPCA. Chemical Engineering Science 59, 5897–5908 (2004)
10. Cui, P., Li, J., Wang, G.: Improved kernel principal component analysis for fault detection. Expert Systems with Application 34, 1210–1219 (2008)
11. Baudat, G., Anouar, F.: Kernel-based methods and function approximation. In: Proceedings of international conference on neural networks, Washington, DC, pp. 1244–1249 (2001)
12. Wu, Y., Huang, T.S., Toyama, K.: Self-supervised learning for object based on kernel discriminant-EM algorithm. In: Proceedings of international conference on computer vision, Kauai, HI, vol. 1, pp. 275–280 (2001)

Machinery Vibration Signals Analysis and Monitoring for Fault Diagnosis and Process Control

Juan Dai[1], C.L. Philip Chen[2], Xiao-Yan Xu[3], Ying Huang[4], Peng Hu[5], Chi-Ping Hu[1], and Tao Wu[1]

[1] The Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming, 650091 China
jdkust@yahoo.com.cn
[2] The Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX 78249 USA
[3] The Electrical Engineering Department, Shanghai Maritime University, Shanghai, 200135 China
[4] Yunnan University of Traditional Chinese Madicine, Kunming, 650021 China
[5] The Faculty of Applied Technology, Kunming University of Science and Technology, Kunming, 650091 China

Abstract. The vibration signals contain a wealth of complex information that characterizes the dynamic behavior of the machinery. Monitoring rotating machinery is often accomplished with the aid of vibration sensors. Transforming this information into useful knowledge about the health of the machine can be challenging due to the presence of extraneous noise sources and variations in the vibration signal itself. This paper describes applying vibration theory to detect machinery fault, via the measurement of vibration and voice monitoring machinery working condition, also proposes a useful way of vibration analysis and source identification in complex machinery. An actual experiment case study has been conducted on a mill machine. The experiment results indicate that fewer sensors and less measurement and analysis time can achieve condition monitoring, fault diagnosis, and damage forecasting. Further applications allow feedback to the process control on production line.

Keywords: Reliability, fault detection, condition monitoring, vibration analysis, process control.

1 Introduction

A developing fault in a machine or a structural component will always show up as an increasing vibration at a frequency associated with the fault; however the fault might be well developed before it affects either the overall vibration level or the peak level in the time. Noise signals measured at regions in proximity to, and vibration signals measured on the external surfaces of machines contain vital information about a machine's running condition [1.5]. When machines are in a good condition, their noise and vibration frequency spectra have characteristic shops. As faults begin to develop,

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 696–703, 2008.
© Springer-Verlag Berlin Heidelberg 2008

the frequency spectra change. This is the fundamental basis for using noise and vibration measurement and analysis in condition monitoring [2]. A frequency analysis of the vibration will give a much earlier warning of the fault, since it is selective, and will allow the increasing vibration at the frequency associated with the fault to be identified. Machine condition, machine faults and on-going damage can be identified in operating machines by fault symptoms [3.4]. Therefore vibration analysis can identify developing problem before they become too serious and cause breakage, using vibration and noise analysis, the condition of a machine can be constantly monitored and more detailed analyses can be made to determine the health of a machine and identify any faults that may be arising or that already exist.

Comprehensive equipment condition-monitoring program is essential for ensuring maximum utility of assets in most industries. To effectively monitor large numbers of assets of various types, a scalable, flexible and usable approach to condition monitoring is warranted. It is becoming increasingly apparent that condition monitoring of machinery reduces operational and maintenance cost and provides a significant improvement in plant availability. The obtained complex information from the measurement signals has to be reduced to the trend of few characteristic values to forecast the development of damages in the near future respectively to allow feedback to the process control.

2 Noise and Vibration Analysis Detected Typical Faults

Common machinery consists of a driver or a prime mover, such as an electric motor. Other prime movers include diesel engineers, gas engineers, steam turbines and gas turbines. The driven equipment could be pumps, compressors, mixers, agitators, fans, blowers and others, at time when the driver equipment has to be driven at speeds other than the prime mover, a gearbox or a belt is used. Each of these rotating parts is further comprised or simple components such as stator, rotors, seals, bearings and so on. When these components operate continuously at high speeds, wear and failure is imminent [7]. When defects are developed in these components, they give rise to higher vibration levels. It can be stated that whenever either one or more parts are unbalanced, misaligned, loose, eccentric, out of tolerance dimensionally, damaged or reacting to some external force, higher vibration levels will occur. The common noise and vibration sources include mechanical noise, electrical noise, aerodynamic, and impulsive noise. Mechanical noise is associated with fan/motor unbalance, bearing noise, structural vibration, reciprocating forces, etc. electrical noise is generally due to unbalance magnetic forces associated with flux density variations and air geometry, brush noise, electrical arcing, etc. Aerodynamic noise is related to vortex shedding, turbulence, acoustic modes inside ducts, pressure pulsations, etc. Finally, impact noise is generated by sharp, short, forceful contact between two or more bodies. Vibration monitor and analysis aim to correlate the vibration response the system with specific defects that occur in the machinery and its components, trains or even in its mechanical structures. Some of the more common faults or defects that can be detected using vibration or noise analysis are summarized in Table 1.

Table 1. Typical Faults Can Be Detected With Noise and Vibration Analysis

Item	Fault
Gears	tooth meshing faults, Eccentric gears
	Cracked and/or worn teeth
Rotors and shafts	Unbalance, Bent shafts
	Eccentric journals, Loose components
	Critical Speeds, Cracked shafts
Rolling element bearings	Pitting of race and ball/roller
	Other Rolling Element
	Journal/bearing rub
Journal Bearing	Defects, Oil whirl
	Oval or barreled journals
Flexible couplings	Misalignment, Unbalance
Electrical machines	Unbalanced magnetic pulls
	Broken/damaged rotor bars
	Air gap geometer variations
Miscellaneous	Structural and foundation faults

3 An Important Vibration Signal Process FFT Analysis

Numerous analysis techniques are available for condition monitoring of machinery or structural components with noise vibration signals. The commonly used signal analysis techniques are magnitude analysis, time domain analysis and frequency analysis. State of technology in vibration monitoring of rotating machines is related to the calculation of standard deviation and/or maximum values, their comparison with thresholds and their trend behavior to determine increased wear or changes in the operation conditions. Time domain averaging is applied to separate speed related information from superimposed resonances and stochastically excitations like mechanical or flow friction [6]. Spectrum analysis with special phase constant averaging routines allows determining machine specific signatures by magnitude and phasing relation.

The measured vibrations and voices are always in analog form (time domain), however the time domain single is difficult to condition monitoring of machinery. Firstly, individual contributions from components in a machine to the overall machine vibration and noise radiation are generally very difficult to identify in the time domain, especially if there are many frequency components involved. This becomes much easier in the frequency domain, since the frequencies of the major peaks can be readily associated with parameters such as rotation frequencies. Secondly, a developing fault in a machine will always show up as an increasing vibration at a frequency associated with the fault. However, the fault might be well developed before it affects either the overall vibration level or the peak level in the fault in the time domain. A frequency analysis of the vibration will give a much earlier warning of the fault, since it is selective, and will allow the increasing vibration at the frequency associated with the fault to be identified. When using vibration and voice as a diagnostic tool, the measured analog signal needs to be transformed to frequency domain. As shown in Figure 1, the fast Fourier transform (FFT) is widely used both by commercially available spectrum analysis and by computer based signal analysis systems [1].

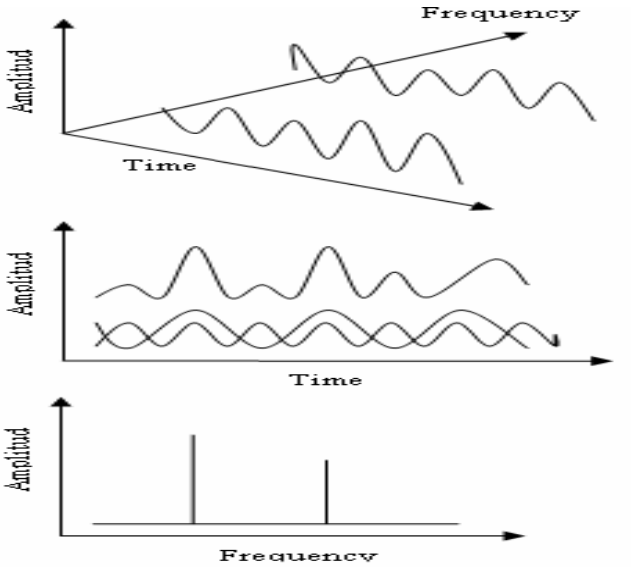


Fig. 1. Fourier Transform (Schematic illustration of time and frequency components)

A general Fourier Transform pair, X (ω) and x (t), is

$$X(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt \cdot \tag{1}$$

And,

$$x(t) = \int_{-\infty}^{\infty} X(\omega)e^{i\omega t}d\omega \cdot \tag{2}$$

Because classical Fourier theory is only valid for functions which are absolutely integrable and decay to zero, the transform X (ω) will only exist for a random signal which is restricted by a finite time interval. Thus, the concept of a finite Fourier transform, X (ω, T) is introduced. The finite transform of a time signal χ (t) is given by,

$$F\{x(t)\} = X(\omega, T) = \frac{1}{2\pi}\int_{0}^{T} x(t)e^{-i\omega t}dt \cdot \tag{3}$$

F{x (t) } represents a forward Fourier transform, and F^{-1} {x (t) } represents an inverse Fourier transform [5].

4 The Complex Machinery's Vibration Monitoring System

Vibration or noise analysis in plant often is involved complex machinery, and a considerable amount of deterioration or damage may occur in there. Therefore vibration source identification and ranking in complex machinery is very important for the implementation condition monitoring and fault detection. The normal techniques use a lot of sensors, which consumes in life in every operation. The paper discusses a new useful way of vibration analysis and source identification in complex machinery based on advanced signal processing techniques. The vibration monitoring and diagnostic method uses three of measurement intensity measurement techniques to

identify vibration source in complex machinery shown in Figure 2. The vibration monitoring system only uses 6 sensors (3 microphones and 3 accelerometers). Among them, 2 microphones are used to measure sound intensity, one microphone and one accelerometer are used to measure surface intensity and 2 accelerometers are used to measure vibration intensity. Figure 2 shows a vibration monitoring and fault diagnosis system for complex machinery. In practice, care has got to be exercised with all intensity techniques and several procedures have been developed to reduce any phase errors between the two microphones. These include (1) the usage of phase matched microphones, (2) microphone switching procedures, which, in principle, eliminate the phase mismatch, and (3) a frequency response function procedure which corrects the intensity data with the measured phase mismatch.

With the surface intensity technique, an accelerometer is mounted on the surface of the vibration structure and a pressure microphone is held in close proximity to it. It is assumed that the velocity of the vibration surface is equal to the acoustic particle velocity and that the magnitude of the pressure does not vary significantly from the vibrating surface to the microphone [8]. The resultant surface intensity normal to the surface (sound intensity at the surface of the structure) is given as

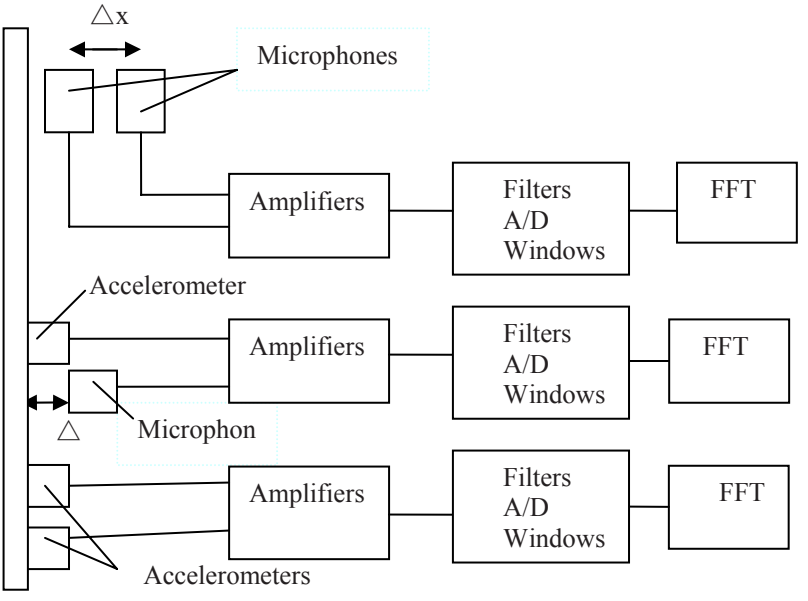


Fig. 2. Vibration monitoring and fault diagnosis for complex machinery

$$Ix = \frac{1}{2\Pi}\int_0^f \frac{Q_{ap}\cos\phi + C_{PA}\sin\phi}{f}df \ , \tag{4}$$

where Q_{pa} the quadrature spectrum (imaginary part of the one-sided cross-spectral density), C_{pa} is the coincident spectrum (real part of the one-sided cross-spectral density), and $\oint$ is the phase shift between the two signals due to the instrumentation

and due to the separation distance between the microphone and the accelerometer. This phase shift has to be accounted for in the analysis. The time lag phase shift can be evaluated for each measurement point and the instrumentation phase shift has to be evaluated during the calibration procedure.

$$\Pi = \sum_{j=1}^{n} \left\{ \sum_{i=1}^{n} \left(Q_{pa} \cos \phi_j + C_{pa_{ij}} \sin \phi_j \right) A_i \right\} \frac{\Delta f}{2\pi f_j}, \tag{5}$$

where n is the number of data points in the frequency domain, N is the number of area increments (A_i)on the whole surface, and $\triangle f$ is the frequency resolution (i. e. $\triangle f = f/n$). The main advantage of the surface intensity is that information is not required about the radiation ratio of a vibrating surface. It is also useful in highly reverberant spaces where a reverberant field exists very close to the surface of a machine. It is main disadvantage is that the phase difference between the microphone and the accelerometer has to be accurately accounted for.

The vibration intensity measurement technique is used to identify free-field energy flow due to bending waves in a solid body. It can be shown that the vibration intensity, I_v, in a given direction is

$$I_v = \frac{\left(B_{P_s}\right)^{1/2}}{2\pi f \Delta x} \int_0^T \left\{ \frac{(a_1 + a_2)}{2} \int (a_2 - a_1) d\tau \right\} dt , \tag{6}$$

where B is the bending stiffness of the structure, P_s is the mass per unit area, $\triangle x$ is the separation between the two accelerometers, and a_1 and a_2 are the two accelerometer signals. It is noted that the scaling factor is frequency dependent.

5 Condition Monitoring and Fault Detection on Mill Machine

The vibration monitoring method designed in this paper, are applied on a mill machine. The mill machine and its vibration condition monitor system are shown in Figure 2. In this system, 3 accelerometers and 3 microphones are mounted in close critical areas. In the actual experiment, continuously monitors signals come from the accelerometers and microphones. High pass filtering increases significantly the probability of fault detection, too. These piezoelectric sensors generate analog signals that measure the component vibration. The analyses system uses National Instruments Lab VIEW developing a PC-based monitoring system with accelerometers and microphones to predict failure of the complex mill machine components. In this case, sound intensity and vibration intensity calculate was applied to analysis where the signal consists of the general descriptive values of the time information (variance, kurtosis) in certain specific frequency ranges as well as peak analysis known from the evaluation of signals. A reference normal operation of all values are normalized to “one”. The actual signatures can then be calculated as differences for comparison purposes. The difference to the reference frequency spectra then is detected with a developing fault. Figure 3 shows the measurement result with sound intensity measurement with vibration fault source. For comparison purposes, of different operation conditions, the actual signals frequency, amplitude and the difference to the reference normal running signals are calculated. The signal classes can be separated depending on the

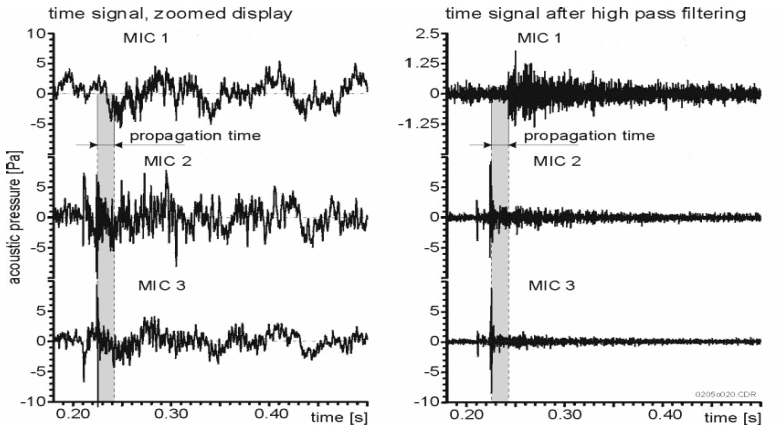


Fig. 3. Determination of failure source using intensity measurement analysis

running condition. The vibration condition monitoring system of the mill machines can select various levels of response when registering anomalies or impending failure. Figure 4 is a typical frequency spectrum vibration signals with different degrees of bearing housing looseness on the mill machine. This setup can alert the maintenance supervisor immediately for diagnosis or repair. The visualized signatures in time and frequency domains have to be summarized to obtain an actual condition classifiers of normal operation to get an automatically feedback for process regulation.

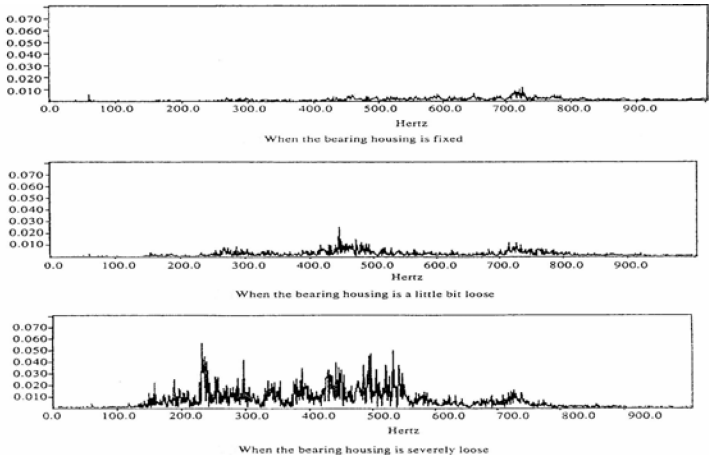


Fig. 4. Frequency spectra under different degrees of bearing housing looseness on mill machine

6 Conclusion

The experiment results show that vibration, sound and acoustic emission combined with intensity technique like sound intensity and vibration intensity measurement

techniques are more convenient and reliable for condition monitoring in complex machinery and quality control than most of the standard methods, like heat, current, force and normal vibration measurements as well as power consumption, used in commercially available systems. The proposed monitoring system presents using fewer sensors and takes less measure and analysis time. The obtained information from the measurement signals is able to forecast the development of damage. The proposed techniques have a great potential to improve industrial production lines utilization rate and product quality by machines condition monitoring. As a result, lower in running operation and maintenance costs and increased in productivity and efficiency can be achieved, further application allow feedback to the process control on production line.

Acknowledgments

This research is supported by the P.R. China Ministry of Education and the authors acknowledge financial support The China Scholarship Council (C.S.C).

References

1. Barber, A.: Handbook of Noise and Vibration Control, 6th edn. Elsevier Advanced Technology Publications, UK (1992)
2. Dai, J., Li, G.Y.: Vibration Test Analysis and Fault Diagnosis on Three Press Roll Calendar. Journal of Kunming University of Science and technology (Periodical style) 4, 68–73 (1997)
3. Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: Grid Information Services for Distributed Resource Sharing. In: 10th IEEE International Symposium on High Performance Distributed Computing, pp. 181–184. IEEE Press, New York (2001)
4. Norton, M.P., Denis, K.: Fundamentals of Noise and Vibration Analysis for Engineers, Cambridge, UK (2003)
5. Liao, B.Y.: Basics of Machinery Fault Detection. Metallurgical Industry Press, Beijing (1999)
6. Reimche, W., Südmersen, U.: Basics of Vibration Monitoring for Fault Detection and Process Control, vol. 6, pp. 2–6. Rio de Janeiro, Brazil (2003)
7. Gade, S.: Sound Intensity and its Application in Noise Control. Sound and Vibration 3(85), 14–26 (1985)
8. Schaffer, C., Girdhar, P.: Machinery Vibration Analysis & Predictive Maintenance. Elsevier Press, Netherlands (2004)

An Ultrasonic Signal Processing Technique for Extraction of Arrival Time from Lamb Waveforms*

Haiyan Zhang[1], Xiuli Sun[1], Shixuan Fan[1], Xuerui Qi[2], Xiao Liu[2], and Donghui Lu[2]

[1] School of Communication and Information Engineering, The Key Lab of Specialty Fiber Optics and Optical Access Network, Shanghai University, Shanghai 200072, China
[2] School of Electronic, Information and Electrical Engineering, Shanghai Jiaotong University, Shanghai 200030, China
hyzh@shu.edu.cn, sxl.xd@163.com, qiuqia17@126.com, qixuerui1998@hotmail.com, 0813snake@163.com, dhlu@shu.edu.cn

Abstract. Travel time Lamb wave tomography has been shown to be an effective nondestructive evaluation (NDE) technique for plate-like structures. The methods used previously to extract the arrival times of the fastest or multi-mode Lamb wave modes are mostly based on enveloping techniques, and various time-frequency methods such as Wigner-Ville distributions, short-time Fourier transforms, and recently explored wavelet transform (WT). Frankly speaking, the uses of signal processing methods improve the accuracy of the arrival time extraction to a great extent relative to interpret the Lamb waveforms directly. Hilbert-Huang transform (HHT) is also an efficient way for analyzing and processing non-stationary signals. The resolving power of time and frequency is restricted from Heisenberg principle in wavelet analysis, while in HHT time resolving power is precise and steady, and frequency resolving power is adaptive according to signal intrinsic characteristics. Conclusion can be made that the HHT method is more adaptive than WT analysis in analyzing non-stationary signals. Based on the above analysis, HHT is attempted to extract multi-mode arrival times from Lamb waveforms in this paper. Comparison with the wavelet transform shows the new method offers higher temporal resolutions in extracting arrival times of multiple Lamb wave modes.

Keywords: Lamb waves; tomography; Hilbert-Huang transform; NDE.

1 Introduction

Ultrasonic Lamb wave tomography has been receiving considerable attention in recent years and traveltime tomography is a subject of intensive study nowadays. Traveltime tomography is a method of estimating object velocity structure for a given set of arrival times. Accurate measurement of the arrival times of Lamb wave modes is a critical step in the tomographic imaging process because it strongly affects all subsequent steps. Different signal processing techniques have been developed in either the time-domain or the time-frequency domain[1-2]. The previous work to extract the arrival

* Supported by the National Natural Science Foundation of China (10504020).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 704–711, 2008.
© Springer-Verlag Berlin Heidelberg 2008

times of the fastest Lamb wave modes is based on enveloping the signal[3-4]. It is a straightforward and computationally inexpensive method, which has been shown to perform better than other more complicated and computationally expensive time-frequency methods[2]. However, it is incapable of providing arrival time information for multiple modes. In work recently done by Hou et al.[1] and Leonard et al.[2], dynamic wavelet fingerprint (DWFP) and the discrete wavelet transform (DWT) were explored to extract multi-mode Lamb wave arrival time information. Wavelets have become popular in many different types of signals and image processing techniques, particularly advantageous for both the denoising and compression of signals.

Hilbert-Huang Transformation (HHT) is a recently developed method for non-stationary signal processing[5]. By performing empirical mode decomposition (EMD), various frequency components can be effectively separated as intrinsic mode function (IMF). After Hilbert transformation to the IMF sequence, the 3-D discrete time-frequency spectrum containing time, frequency and amplitude can be obtained, which provides very clear time-frequency characteristics with local details. As a good tool of describing signals with non-stable variations, HHT has been used in many areas, such as extraction of seismic precursory information from gravity tide[6], financial time series analysis[7] and so on. In the present application, HHT is used to extract arrival time information for multiple Lamb wave modes.

2 Experiment Setup

Fig.1 shows the schematic diagrams of the experimental setup. A pulser / receiver unit (Panametrics 5900PR, Waltham, MA) was used to excite transducer. An angle transducer with 0.5MHz center frequency was used to generate ultrasonic and a same transducer to receive signals. The received signals were amplified and digitized to 8 bits (HP 54642A, Hewlett-Packard, Palo Alto, CA) at 10 mega-samples/s. Each 2000-point waveform was averaged 64 times in the time domain and then stored on a computer for offline analysis using MATLAB® software. The whole sampling process is monitoring by PC. The thickness of the aluminium plate is 2.5mm. The measurements were made below a frequency-thickness of 1.5 MHz·mm. There are only two modes in this frequency-thickness region, which are A_0 and S_0 modes.

Since many individual measurements are required to develop the projection data in tomographic reconstructions, various measurement schemes, for example, parallel projection, fanbeam tomography, and crosshole geometries had been used and compared in Lamb wave scans. A conclusion is that the crosshole scans from the seismological literature are much better suited to Lamb wave NDE applications than are the parallel and fanbeam scans from the medical imaging literature[3]. Laboratory measurements discussed here use the crosshole geometries, which is a geophysical reconstruction method but also be adapted for use with Lamb wave applications. Fig.2 illustrates the basic setup for crosshole sampling, in which the small circles in two edges represent the different transducer positions, and the big circle in the middle represents the flaw. For each projection the transmitting transducer steps along the left edge as the receiving transducer steps through all the positions on the right edge. For example, in one of the crosshole projections shown in Fig.2 the transmitting transducer is fixed, the receiving transducer steps along the right edges from the upper to

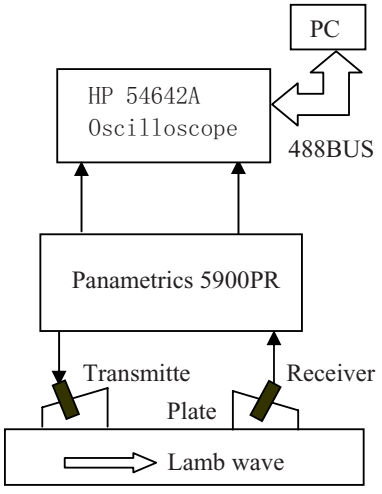


Fig. 1. Experimental setup

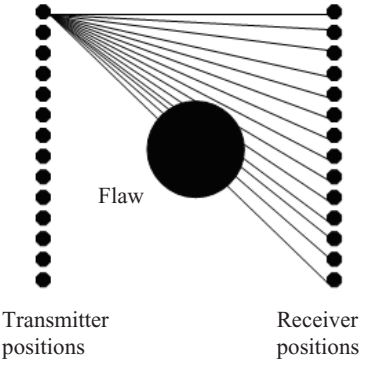


Fig. 2. Crosshole sampling with multiple-transmitter and receiver positions

the lower. After these measurements the first set of crosshole projections is completed. The transmitter is then stepped into the next position. Repeating the above process, the measurements complete the second set of projections.

3 HHT Theory

HHT method is composed of empirical mode decomposition and Hilbert transform. First, the original signal is decomposed into a group of IMF, and then Hilbert transform is performed to each IMF component. Actual signals are usually complicated, which may not always satisfy IMF condition. So, Huang et al[5] put forward a creative hypothesis: Any signal is composed of different IMF; a complicated signal can contain several IMF signals at any time; if superposition is found between the modes, then compound signals are formed. On this basis, Huang further pointed out that EMD can be used to filtrate out IMF. After filtration the original signal can be expressed by the sum of finite IMF and a residual component

$$x(t) = \sum_{j=1}^{n} c_j(t) + r_n(t) \tag{1}$$

Thus, a decomposition of the complicated signal into n-empirical modes c_j's is achieved, and a residue r_n obtained can be either the mean trend or a constant. Having obtained the intrinsic mode function components, one will have no difficulty in applying the Hilbert transform (HT) to each IMF component.

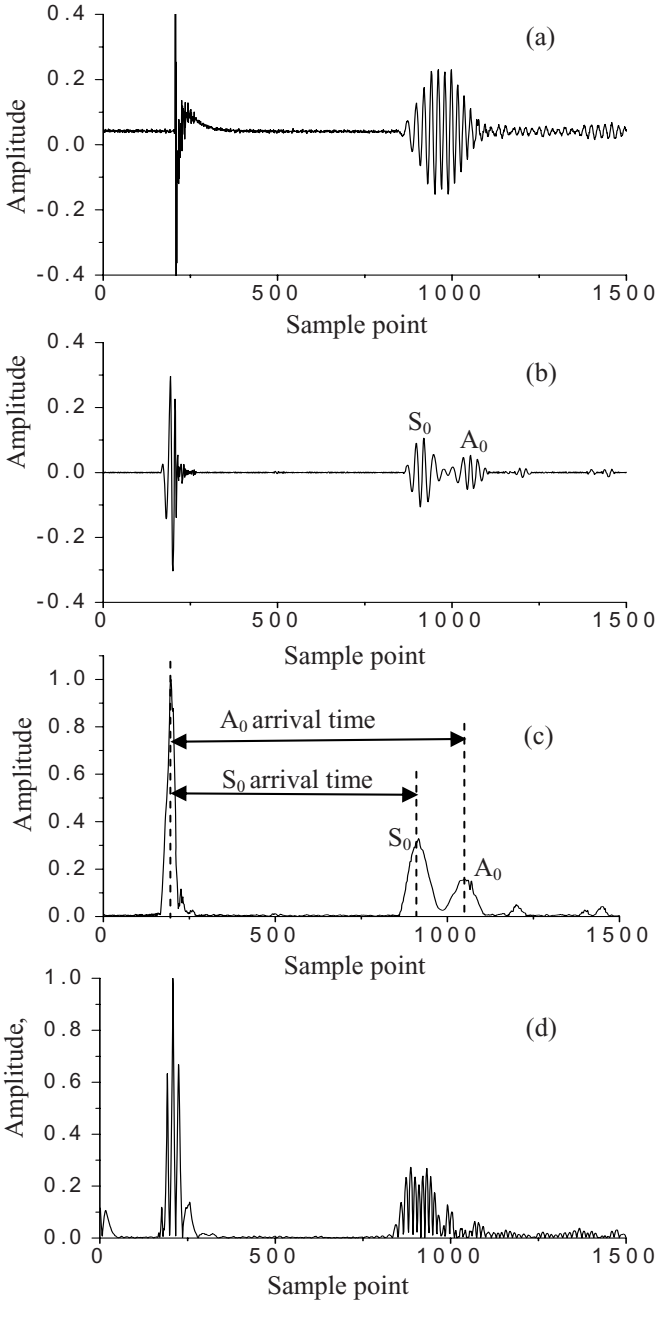


Fig. 3. Illustration of the HHT-based technique to measure the arrival times of multiple Lamb wave modes. (a) Original waveform; (b) An IMF component after EMD decomposition; (c) HT to IMF component in (b); (d) Wavelet envelope.

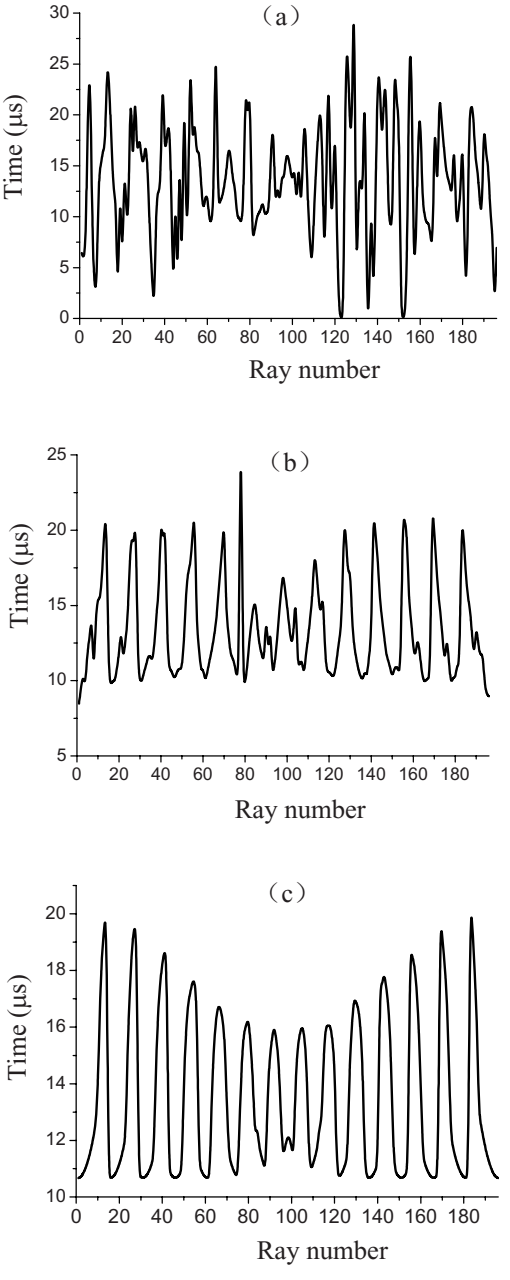


Fig. 4. The experimentally estimated and theoretical arrival times of the S0 mode for an aluminium plate with a 4-mm-diameter circular hole. (a) Wavelet enveloping technique; (b) HHT method; (c) Theoretical simulation.

4 Arrival Time Extraction

Fig.3 (a)~(c) illustrates the HHT-based technique to measure the arrival times of multiple Lamb wave modes. As comparison, the wavelet envelope technique is also given in Fig.3(d). The x-axis in each of these samples corresponds to arrival time. The y-axis in each is amplitude in arbitrary units. It is obvious in Fig.3(b) that two modes, i.e. S_0 and A_0 modes are included in the IMF component. Performing Hilbert transform (HT) to the IMF component in Fig.3(b), the arrival times of two modes can be extracted accurately, as shown in Fig.3(c), where the time difference between starting pulse and each peaks represents the arrival time of each mode. In Fig.3(d), the wavelet envelope reveals many "false peaks", which cannot give correct arrival time sorting for multi-mode Lamb waves.

For current scanning system shown in Fig.2, a total of 14×14=196 waveforms resulting from all possible transmitter-receiver positions are recorded according to the crosshole geometry. The shortest distance between transmitter and receiver lines was 81mm. The experimentally estimated and theoretical arrival times of the S_0 mode for an aluminium plate with a 4-mm-diameter circular hole are shown in Fig.4. The x-ray for each is ray number, and the y-axis for each is arrival time in μs. The overall oscillating nature for these plots is due to the changing ray length inherent in the criss-cross geometry of the tomography technique. Visual inspection reveals that the arrival times extracted by HHT method are distributed rather smoothly, and agree very well with theoretical values.

For comparison, the deviations of the arrival times between different extraction methods from measured Lamb waveforms and theoretical simulation are shown in Fig.5. It can be seen that the time deviations between HHT method and theoretical simulation (dotted line) are much smaller than the deviations between wavelet enveloping technique and theoretical simulation (solid line), which verifies the performance of the proposed algorithm.

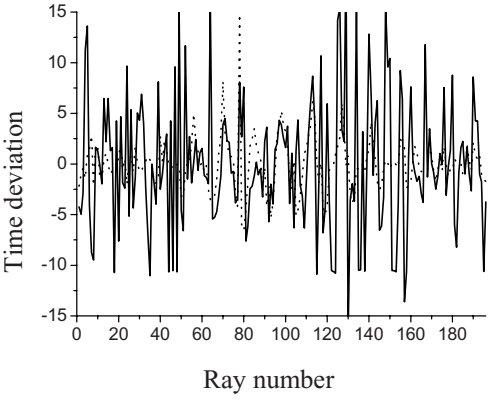


Fig. 5. The deviations of the arrival times between different extraction methods from measured Lamb waveforms and theoretical simulation. The solid line represents the deviations of arrival times between the wavelet enveloping technique and theoretical simulation. The dotted line represents the deviations of arrival times between the HHT method and theoretical simulation.

Crosshole tomographic reconstructions using the simultaneous iterative reconstruction technique (SIRT) for different arrival time extraction techniques are shown in Fig.6. The detected flaw is a flat bottom hole. Compared with the reconstruction result using the simulated data, it can be seen that the flaw is more accurately sized using the arrival time data of HHT than that of the wavelet enveloping.

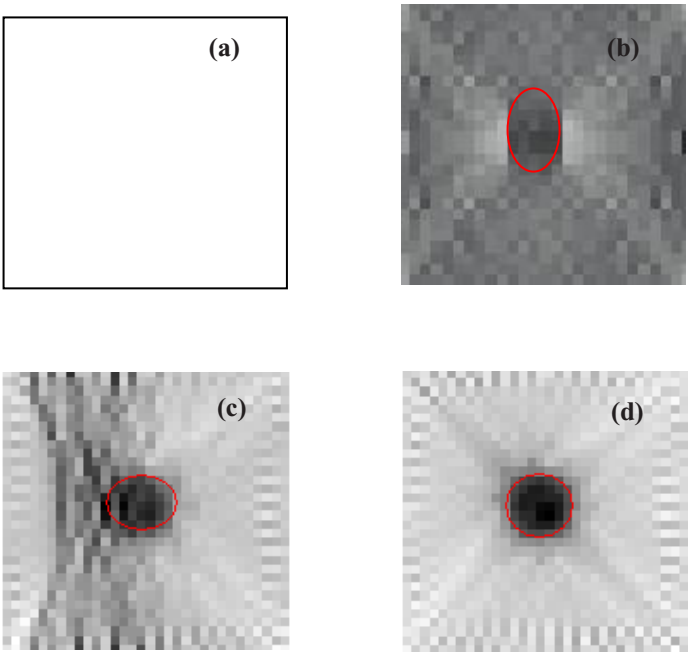


Fig. 6. Tomographic reconstructions using different arrival time extraction techniques: (a) Cartoon of the plate; (b) wavelet enveloping; (c) HHT; (d) simulation

5 Conclusions

This paper presents the application of the HHT technique to extracting arrival times of multiple Lamb wave modes. Although Lamb waveforms are quite complex, the application of HHT technique leads to automatic detection of multiple Lamb wave modes in a systematic way. A comparison of the HHT and wavelet transform algorithms to theoretical data shows that the new technique has the capability to extract more accurate arrival times, and has the potential to lead to even better tomographic reconstructions that accurately locate and size flaws in plates.

Only a single mode is used in this research although HHT method has the capability to sort multiple Lamb wave modes. Since each mode has different through-thickness displacement values, for a specific flaw present in the plate, such as surface corrosion, some modes may be more sensitive than other modes. Therefore, to sufficiently exert the performance of the suggested algorithm, further work is to exploit several Lamb wave modes and generate corresponding tomographic images.

References

1. Hou, J., Leonard, K.R., Hinders, M.K.: Automatic multi-mode Lamb wave arrival time extraction for improved tomographic reconstruction. Inver. Prob. 20, 1873–1888 (2004)
2. Leonard, K.R., Hinders, M.K.: Multi-mode Lamb wave tomography with arrival time sorting. J. Acoust. Soc. Am. 117(4), 2028–2038 (2005)
3. Malyarenko, E.V., Hinders, M.K.: Ultrasonic Lamb wave diffraction tomography. Ultrasonics 39(4), 269–281 (2001)
4. Leonard, K.R., Malyarenko, E.V., Hinders, M.K.: Ultrasonic Lamb tomography. Inver. Prob. 18, 1795–1808 (2002)
5. Huang, N.E., Shen, Z., Long, S.R.: The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. A (1998)
6. Zhang, L., Fu, R.S., Zhou, Z., et al.: Extraction of seismic precursory information from gravity tide at Kunming station based on HHT. ACTA Seis. Sinica 20(2), 234–238 (2007)
7. Huang, N.E., Wu, M.L., Qu, W.D., et al.: Applications of Hilbert-Huang transform to non-stationary financial time series analysis. Applied Stochastic Model in Business and Industry 19(3), 245–268 (2003)

Application Research of Support Vector Machines in Dynamical System State Forecasting

Guangrui Wen[1,2], Jianan Yin[2], Xining Zhang[1], and Ying Jin[2]

[1] Research Institute of Diagnostics & Cybernetics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[2] Xi'an Shaangu Power Co., Ltd, Xi'an, Shaanxi 710611, China
{grwen,zhangxining}@mail.xjtu.edu.cn,
{sgyja,jinying_611}@163.com

Abstract. This paper deals with the application of a novel neural network technique, support vector machines (SVMs) and its extension support vector regression (SVR), in state forecasting of dynamical system. The objective of this paper is to examine the feasibility of SVR in state forecasting by comparing it with a traditional BP neural network model. Logistic time series are used as the experiment data sets to validate the performance of SVR model. The experiment results show that SVR model outperforms the BP neural network based on the criteria of normalized mean square error (NMSE). Finally, application results of practical vibration data state forecasting measured from some CO2 compressor company proved that it is advantageous to apply SVR to forecast state time series and it can capture system dynamic behavior quickly, and track system responses accurately.

Keywords: Support vector machines (SVMs), Dynamical System, State Forecasting.

1 Introduction

Recently, support vector machines (SVMs) have been proposed as a novel technique in time series forecasting [1-3]. SVMs are a very specific type of learning algorithms characterized by the capacity control of the decision function, the use of the kernel functions and the sparsity of the solution [4-6]. Established on the unique theory of the structural risk minimization principle to estimate a function by minimizing an upper bound of the generalization error, SVMs are shown to be very resistant to the over-fitting problem, eventually achieving high generalization performance in solving various time series forecasting problems [7-11]. Another key property of SVMs is that training SVMs is equivalent to solving a linearly constrained quadratic programming problem so that the solution of SVMs is always unique and globally optimal, unlike other networks' training which requires non-linear optimization with the danger of getting stuck into local minima.

State forecasting of a dynamic system entails the current and previous conditions to forecast the future states of the dynamic system. In a machine, for example, a reliable prognostic system can predict the fault propagation trends, and provide an accurate

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 712–719, 2008.
© Springer-Verlag Berlin Heidelberg 2008

alarm before a fault reaches critical levels so as to prevent machinery performance degradation, malfunctions, and even catastrophic failures [12]. In practice, the temporal patterns utilized can be obtained from vibration features, acoustic signals, or the debris in the lubrication oil. Vibration-based monitoring, however, is a well-accepted approach due to its ease of measurement and analysis, and accordingly it is used in this study.

Because SVMs is known to generalize well in most cases and adapts at modeling nonlinear functional relationships that are difficult to model with other techniques. Consequently, we propose to apply the idea of support vector Machines for function estimation, i.e. Support Vector Regression (SVR) to build prediction model and investigate state forecasting of dynamical system.

2 Basic Principles of Support Vector Regression

SVMs was introduced by Vapnik in the late 1960s on the foundation of statistical learning theory [13]. It has originally been used for classification purposes but its principle can be extended easily to the task of regression by introducing an alternative loss function. The basic idea of SVR is to map the input data **x** into a higher dimensional feature space F via a nonlinear mapping φ and then a linear regression problem is obtained and solved in this feature space.

Given a training set of l examples $\{(x_1, y_1), (x_2, y_2), \cdots, (x_l, y_l)\} \subset R^n \times R$, where x_i is the input vector of dimension n and y_i is the associated target. We want to estimate the following linear regression:

$$f(x) = (w \cdot x) + b, \qquad w \in R^n, \qquad b \in R \tag{1}$$

Here we consider the special case of SVR problem with Vapnik's ε – insensitive loss function defined as:

$$L_\varepsilon(y, f(x)) = \begin{cases} 0 & |y - f(x)| \le \varepsilon \\ |y - f(x)| - \varepsilon & |y - f(x)| > \varepsilon \end{cases} \tag{2}$$

The best line is defined to be that line which minimizes the following cost function:

$$R(w) = \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{l} L_\varepsilon(f(y_i, x_i)) \tag{3}$$

where C is a constant determining the trade-off between the training errors and the model complexity. By introducing the slack variables ξ_i, ξ_i^*, we can get the equivalent problem of Eq.3. If the observed point is 'above' the tube, ξ_i is the positive difference between the observed value and ε. Similarly, if the observed point is 'below' the tube, ξ_i^* is the negative difference between the observed value and $-\varepsilon$. Written as a constrained optimization problem, it amounts to minimizing:

$$\frac{1}{2}\| w \|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{4}$$

subject to:

$$\begin{aligned} & y_i - (w \cdot x_i) - b \le \varepsilon + \xi_i \\ & (w \cdot x_i) + b - y_i \le \varepsilon + \xi_i^* \end{aligned} \tag{5}$$

To generalize to non-linear regression, we replace the dot product with a kernel function $K(\cdot)$ which is defined as $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$. By introducing Lagrange multipliers α_i, α_i^*, which are associated with each training vector to cope with both upper and lower accuracy constraints, respectively, we can obtain the dual problem that maximizes the following function:

$$\sum_{i=1}^{l}(\alpha_i - \alpha_i^*)y_i - \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) - \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) \tag{6}$$

subject to:

$$\begin{aligned} & \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \\ & 0 \le \alpha_i, \alpha_i^* \le C \qquad i = 1,2,\cdots,l \end{aligned} \tag{7}$$

Finally, the estimate of the regression function at a given point **x** is then:

$$f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i, x) + b \tag{8}$$

3 Experiment Verification

The experimental verification was conducted by a simulated time series produced by the logistic map. In order to measure the ability of SVR and traditional BP neural network model, the normalized mean square error (NMSE), also called prediction error, has been utilized:

$$E = \frac{1}{2N}\sum_{k=0}^{N-h}(x(k+h+1) - \hat{x}(k+h+1))^2 \tag{9}$$

where h is the prediction length, $x(k+h+1)$, and $\hat{x}(k+h+1)$ are the real and predicted values of the time series at instant $k+h+1$; N is the number of patterns.

3.1 Prediction at the Logistic Map

The simulated time series is given by following equation:

$$x(k+1) = \lambda x(k)(1 - x(k)), \text{ When } \lambda = 3.97 \text{ and } x(0) = 0.5 \tag{10}$$

The equation describes a strong chaotic time serials and we initialize parameter k from $k=0$ to $k=100$. Table 1 and 2 show the predicted errors of two models over the training data and predictied data respectively and Fig.1 displayed the multi-step prediction curves of two models respectively.

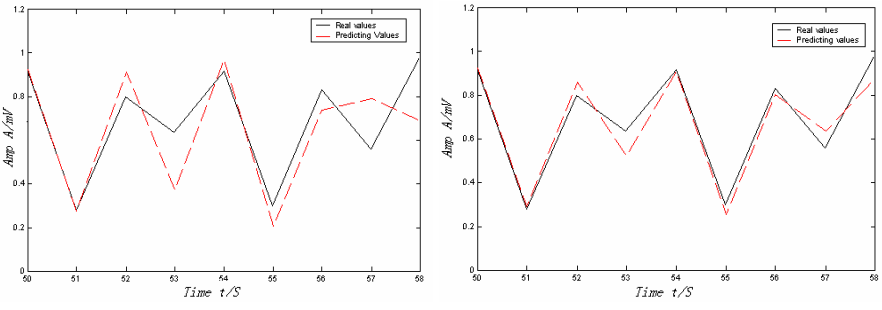
Remark 1. Traditional neural network model (3-10-1) means that the number o*f input layer is 3, the number of hidden layer is 10 and the number of output layer is 1.*

Table 1. Logistic time series: prediction errors over the training data

Length h	Traditional model (3-10-1)	SVR model
0	0.0015	0.0015
4	0.0550	0.0130
7	0.0850	0.0135

Table 2. Logistic time series: prediction errors over the predicting data

Length h	Traditional model (3-10-1)	SVR model
0	0.002	0.002
4	0.072	0.015
7	0.105	0.026



(a) (b)

Fig. 1. The left figure (a) displays the prediction curve of traditional neural networks model using (3-10-1) structure and the right figure (b) shows the prediction result of SVR model, ε is fixed at 0.01 (the constant C=5000)

It is evident that SVR model's prediction accuracy is much higher than that of traditional BP neural networks model.

3.2 Analysis of Experiment Results

Table 1 and Table 2 give the comparative results of two prediction models in 8-step prediction. The experiments show that the errors of SVR model on training set are

less than that of traditional BP neural networks model about 0.07. This indicates that the prediction performance of SVR model is superior to that of traditional BP neural networks model. Fig. 1 gives the prediction results over test set at different models. The figure shows that the errors of two prediction models are very close in 1-6 step prediction. But in 7-8 step prediction process, Fig.1 (b) shows that the SVR model possess the high forecasting ability in long term prediction. Table 2 shows the same results and the errors of SVR model on test set is less than that of traditional BP neural networks model about 0.08. This indicates that the SVR model has high robustness and anti-jamming ability in long-term prediction. So, in present paper, we adopt the SVR model to realize the long-term state forecasting of dynamical system.

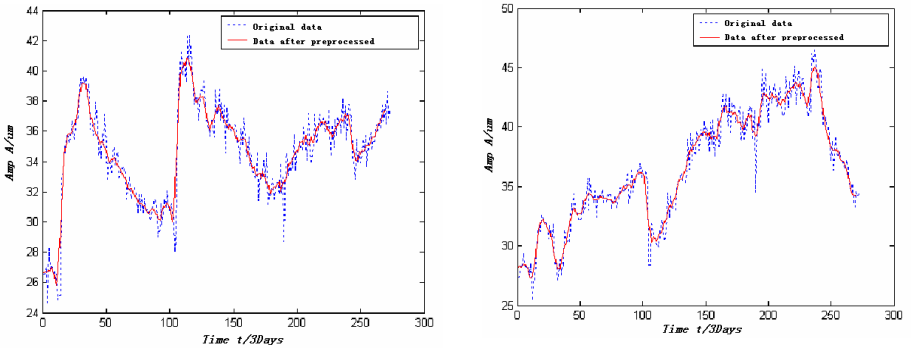
4 Application of Long-Term Forecasting

Trending method is well known in condition monitoring of dynamical system [14-16]. Moreover, in monitoring systems the peak-to-peak values of vibration signal are used as the trending object to supervise if there is any change in vibration condition of machine.

4.1 Preprocessing of Vibration data

In order to reduce the random fluctuation of practical vibration data and improve the prediction precision, local moving average method is applied to preprocess of vibration data. Moreover, the trend component can also be extracted as the evidence of long-term forecasting. Suppose we are given rough time series $\{x_1, x_{2,} \cdots, x_N\}$ and the window length is fixed to *s*, the data after local moving average processing is:

$$x'_i = \frac{1}{s}\sum_{j=i}^{i+s-1} x_j, i = 1, 2, ..., N-s+1 \tag{11}$$



(a) Peak-peak value of high pressure part 1 (b) Peak-peak values of high pressure part 2

Fig. 2. Preprocessing results of two peak-peak values of vibration data coming from a bearing section of some high pressure part

We applied the stated above method to preprocess the practical vibration data, which are measured from some CO2 compressor. The time span is from 1997.03.06 to 1997.10.07 and sampling interval is 3 days. Fig.2 shows the preprocessed results and the solid line displayed the processed data and the dash line shows the rough data. It is obvious that the data after processed become smoothing and trend component become very clear too.

4.2 Long Term Prediction of Machine Condition

In this section, we will use the preprocessed data 1 to test the performance of two prediction models. The number of all data is 266 and first 188 points will be used in training data set and the last 80 points will be used in testing data set. By optimizing the parameters of SVR, we select $C = 3383$, $\varepsilon = 0.0115$, $\sigma = 94.76$. The (5-11-1) structure is adopted in BP prediction model. Fig.3 shows the predicting trend of peak-to-peak value of this machine with day by using two models.

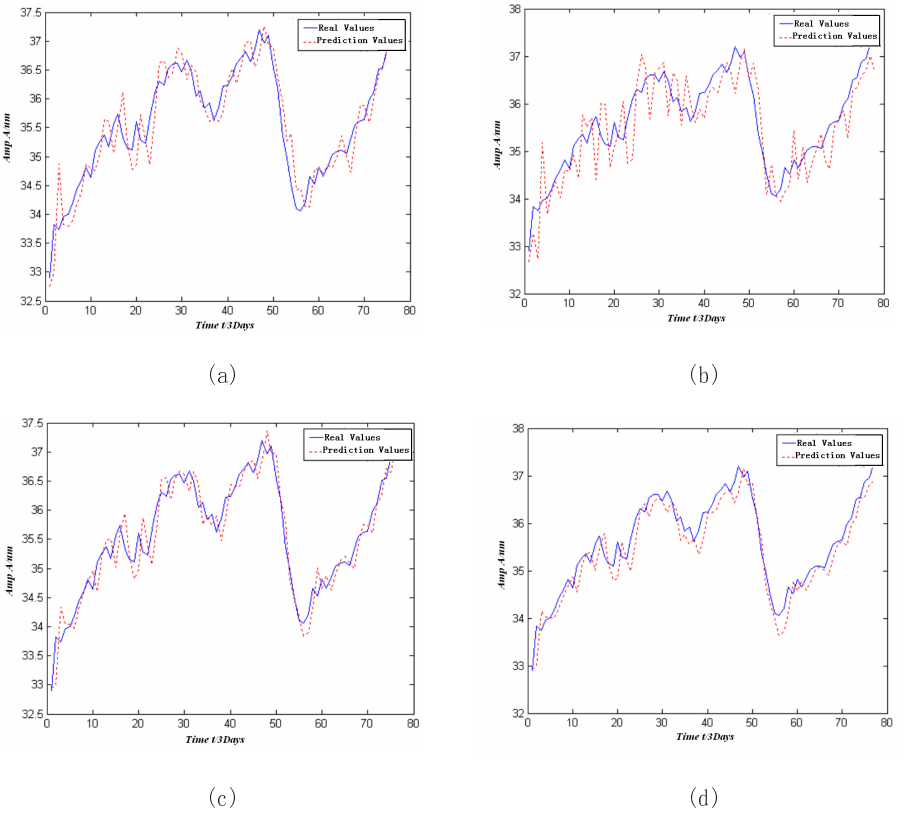


Fig. 3. The forecasting results of two models. (a) Single step prediction results using BP neural networks, (b) two step prediction results using BP neural networks, (c) Single step prediction results using SVR model and (d) two step prediction results using SVR model over testing set of practical vibration data.

Table 3. Prediction errors of two models over the testing data set

Steps	BP	SVR
1	0.04	0.02
2	0.06	0.03

Table.3 shows the prediction errors of two models over the testing data set. The results show that the SVR prediction model possesses a better predictability than traditional BP model in long-term forecasting.

5 Conclusion

A new technique for state forecasting of dynamical system modeling and prediction is proposed in this paper. SVR is adopted to build prediction model. From the experiments and practical application in long-term state forecasting, we can see that the proposed SVR model has a good performance, no matter generalization ability or predictive ability, the SVR is better than traditional BP neural networks model by comparing the prediction performance.

Acknowledgement. This work is supported by Natural Science Foundation of China (No.50775175) and the National High Technology Research and Development Program ("863"Program) of China (No. 2007AA04Z432).

References

1. Mukherjee, S., Osuna, E., Girosi, F.: Nonlinear Prediction of Chaotic Time Series Using Support Vector Machines. In: NNSP 1997: Neural Networks for Signal Processing VII: Proceedings of the IEEE Signal Processing Society Workshop, Amelia Island, FL, USA, pp. 511–520 (1997)
2. Muller, K.R., Smola, A.J., Ratsch, G., Scholkopf, B., Kohlmorgen, J.: Using Support Vector Machines for Time Series Prediction. In: Advances in Kernel Methods—Support Vector Learning, pp. 243–254. MIT Press, Cambridge (1999)
3. Muller, K.R., Smola, J.A., Ratsch, G., Scholkopf, B., Kohlmorgen, J., Vapnik, V.N.: Predicting Time Series with Support Vector Machines. In: Proceedings of the Seventh International Conference on Artificial Neural Networks, Lausanne, Switzerland, pp. 999–1004 (1997)
4. Cristianini, N., Taylor, J.S.: An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, New York (2000)
5. Vapnik, V.N.: An Overview of Statistical Learning Theory. IEEE Trans. Neural Networks 10(5), 988–999 (1999)
6. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer, New York (2000)
7. Thissena, U., van Brakela, R., de Weijerb, A.P., Melssena, W.J., Buydensa, L.M.C.: Using Support Vector Machines for Time Series Prediction. Chemometrics and Intelligent Laboratory Systems 69, 35–49 (2003)
8. Mohandes, M.A., Halawani, T.O., Rehman, S.: Support Vector Machines for Wind Speed Prediction. Renewable Energy 29, 939–947 (2004)

9. Pai, P.F., Hong, W.C.: Forecasting Regional Electricity Load Based on Recurrent Support Vector Machines with Genetic Algorithms. Electric Power Systems Research 74, 417–425 (2005)
10. Gao, L.J.: Support Vector Machines Experts for Time Series Forecasting. Neurocomputing 51, 321–339 (2003)
11. Francis, E., Tay, H., Cao, L.J.: Application of Support Vector Machines in Financial Time Series Forecasting. Omega 29, 309–317 (2001)
12. Wang, W.: An Intelligent System for Dynamical System State Forecasting. In: Wang, J., Liao, X.-F., Yi, Z. (eds.) ISNN 2005. LNCS, vol. 3497, pp. 460–465. Springer, Heidelberg (2005)
13. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer, New York (1995)
14. Yu, H.Y., Bang, S.Y.: An Improved Time Series Prediction by Applying the Layer-by-Layer Learning Method to FIR Neural Networks. Neural Networks, 1717–1729 (1997)
15. Wen, G.R., Qu, L.S.: Multi-step Forecasting Method Based on Recurrent Neural Networks. Journal of Xi'an Jiaotong University, China, 722–726 (2002)
16. Zhang, H.J.: Information Extraction of Mechanical Fault Diagnosis and Predicting. Xi'an Jiaotong University Doctor Thesis, China, pp. 88–102 (2002)

Transformers Fault Diagnosis Based on Support Vector Machines and Dissolved Gas Analysis

Haitao Wu[1], Huilan Jiang[1], and Dapeng Shan[2]

[1] Key Laboratory of Power System Simulation and Control of Ministry of Education, Tianjin University, Tianjin 300072, China
[2] High voltage power supplied company Tianjin 300250 China

Abstract. Transformers are one of the critical equipments of electrical system. This paper analyses the merits and flaws of the existing transformer-fault-diagnosis methods, and then utilizes SVC (Support Vector Classification), one of branches of SVM (Support Vector Machine) which has many advantages like great generalization, effectively avoiding "over-learning" and being able to learn from small scale sample, to analyze transformers dissolved gas. Because of the nonlinear classification relationship between DGA and fault diagnosis, this paper unifies the DGA data, selects proper kernel function, solution, and parameters, formulates the SVC with best effect, sets up a two-floor fault diagnosis tree, and ultimately builds the transformer-fault-diagnosis model. Through the self-test and comparison with Three Ratio Method as well as BP neural network, it is improved that the model built in this paper remarkably enhances the accuracy and efficiency of fault diagnosis.

Keywords: SVM; SVC; DGA; Fault; Diagnosis.

1 Introduction

As one of the critical equipments of electrical system, the transformers' performance directly affects the safe operation and dependability of electrical system. How to realize the diagnosis of transformers promptly and efficiently has been a very important and imperious task. The DGA can take sampling analysis at any time in the run of the transformers, and it can also take an overall analysis by on-line monitoring of the oil chromatographic device. It should be a very perfect monitoring and analysis method in the view of operation. And it has a very important significance to the found of potential fault in the oil filled power equipments. It's very complex to analyze the detailed information of the insulation fault in transformers, and there are three bases: the cumulativeness of the gas production, the rate of the gas production and the characteristics of the gas production under fault [1~3].

The correlative regulation in our country uses the IEC basically which always has the problem of coding missing and coding crossing. So with the advantages of Support Vector Machines such as the learn from small scale sample, global optimization , minimization of the structure risk and the shorter time to diagnosis, this paper erects the fault-diagnosis model using the two layer binary decision tree and deals with the data using reliability data analysis theory. It brings forward the new method which is the diagnosis

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 720 – 727, 2008.
© Springer-Verlag Berlin Heidelberg 2008

of transformers based on Support Vector Machines and dissolved gas analysis to improve the fault diagnosis accuracy and shorten the fault diagnosis time.

2 The Basic Theory of Support Vector Machines

Support Vector Machines theory is a new technique of data-sink. It was brought forward first in the 1990s by Vapnik.

2.1 Mathematics Description of Classification Problem

First of all, we describe the classification problem below with mathematics language.

According to the training set

$$T = \{(x_1, y_1), \cdots, (x_l, y_l)\} \in (\mathbf{x} \times \mathbf{y})^l \tag{1}$$

We will search for a real valued function on the condition of $\mathbf{x} = \mathbf{R}^n$ to use the decision-making function

$$f(x) = \operatorname{sgn}(g(x)) \tag{2}$$

So we can infer the value of y corresponds to any model of x.

2.2 The Methods to Construct the Linear Classification Learning Machine

There are two methods to construct the linear classification learning machine. They are halving the nearest point method and the largest interval method.

2.3 Support Vector Classification

Reference to the language in the field of machine learning, we call the methods to solve the classification problems classification learning machine. When g(x) is a linear function and the classification rule is determined by decision function (2), we call it linear classification learning machine. When g(x) is a nonlinear function, we call it nonlinear classification learning machine.

Generally speaking, classification can be divided into three kinds, linear divisible problems, approximate linear divisible problems and linear indivisible problems. Different kinds of classification learning machines should be used in different kinds of classification problems.

We use a support vector machine as a learning machine. This paper will only introduce the C-support vector classification machine which is also the nonlinear soft-interval classification machine.

It's arithmetic is that:

1. Suppose the known training set

$$T = \{(x_1, y_1), \cdots, (x_l, y_l)\} \in (\mathbf{x} \times \mathbf{y})^l \tag{3}$$

2. Choose the proper kernel function and the proper parameter C, then constructs and solves the optimization problem.

$$\min_{\boldsymbol{\alpha}} \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j \mathbf{K}(x_i \cdot x_j) - \sum_{j=1}^{l} \alpha_j, \tag{4}$$

We can get the optimal solution $\boldsymbol{\alpha}^* = (\alpha_l^*, \cdots, \alpha_l^*)^T$;

3. Choose a positive component of $\boldsymbol{\alpha}^*$, $0 < \boldsymbol{\alpha}^* < C$, and compute the value of b^* according to it

$$b^* = y_j - \sum_{i=1}^{l} y_i \alpha_i^* \mathbf{K}(x_i \cdot x_j); \tag{5}$$

(4) Construct the decision function

$$f(x) = \operatorname{sgn}(\sum_{i=1}^{l} \alpha_i^* y_i \mathbf{K}(x \cdot x_i) + b^*) \tag{6}$$

3 The Fault Diagnosis Based on Dissolved Gas Analysis in Transformers

3.1 The Relation between Transformers' Internal Fault and the Dissolved Gas Content in Oil

According to the phenomena of the fault, the transformers internal fault is mostly the heat fault and the electricity fault. As to the mechanical fault, it will be represented by the heat fault and the electricity fault. The damp problem of transformers is also a latent fault. We must detect it early, or it will turn to be an electricity fault and even a electricity accident.

3.2 The Steps and Methods of Fault Diagnosis

When taking an internal fault diagnosis according to the results of analysis, we should follow the idea below to diagnose the fault:

1. Determine whether there is a fault ;
2. Determine the style of the fault, such as overheating, arc, arc-discharge and partial charge;
3. Diagnose the status of the fault, such as the temperature of hotspot, the power of fault, severity degree, and development trend;
4. Bring forward relevant accident countermeasure, such as whether it can continue to run, the technical safety precautions and scrutiny measure when it continue to run and whether it needs a interior repair [4].

This paper is mainly the second step of the work.

Table 1. Dissolve gas analysis

status	Characteristic failure	C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6
PD	Partial discharge	NS 1)	<0.1	<0.2
D1	Low energy partial discharge	>1	0.1～0.5	>1
D2	High energy partial discharge	0.6～2.5	0.1～1	>2
T1	Heat failure<300° C	NS 1)	>1但NS>1	<1
T2	Heat failure 300° C<t<700° C	<0.1	>1	1～4
T3	Heat failure t>700° C	<0.2 2)	>1	>4

Table 2. The methods to diagnose the failure styles

Coding combination			failure styles
C_2H_2/C_2H_4	CH_4/H_2	C_2H_4/C_2H_6	
0	0	1	Low temperature overheating（<150° C）
	2	0	Low temperature overheating(150° C～150° C)
	2	1	Middle temperature overheating(300° C～700° C)
	0，1，2	2	High temperature overheating（>700° C）
	1	0	Partial discharge
1	0，1	0，1，2	Low energy discharged
	2	0，1，2	Low energy discharged and overheating
2	0，1	0，1，2	Arc-discharge
	2	0，1，2	Arc-discharge and overheating

The methods to diagnose the fault style are characteristic gas method, gas graphic method, two ratio method, triangle graphic method and three ratio method. These days, the criterion of transformers DGA analysis in our country is《analysis and judgment guide of dissolve gas in transformer oil》. It applies the characteristic gas method and three ratio method. Specific judgment criterion is shown in the following tables.

4 The SVC in the Transformer Fault Diagnosis

4.1 Data Processing

The data is the basic of this paper's study. The transformers' fault should be reflected exactly and roundly.

In the actual analysis of the transformers oil chromatography data, and there is prodigious difference between the sensitivity different gas reflects. This paper uses the concept of cumulative frequency in data reliability analysis to unify the transformer oil chromatography data .The target is to reduce the difference between the data while maintaining the integrity of transformer oil chromatography data.

4.2 Model of Fault Diagnosis

The results of once classification to the fault sample of transformer oil chromatographic data are often difficult to meet the requirement. So this paper constructs three classification machines to form a two layer binary decision tree.

The classification machine model is showmen in Fig.1:

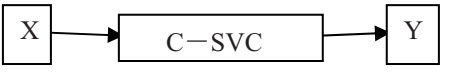


Fig. 1. Classification machine model

This paper uses a Gauss radial basis function as a kernel function,

$$\mathrm{K}(\mathbf{x},\mathbf{x}') = \exp\left(-\frac{\|\mathbf{x}-\mathbf{x}'\|^2}{\sigma^2}\right) \tag{7}$$

This paper uses LIMSVM based on the SMO algorithmic to beg α_i and α_i^* ($i=1,2,\cdots,\mathrm{l}$).We can get the optimal solution $\left(\boldsymbol{\alpha},\boldsymbol{\alpha}^*\right)=\left(\alpha_1,\alpha_1^*,\cdots\alpha_\mathrm{l},\alpha_\mathrm{l}^*\right)^\mathrm{T}$ .Then we compute the value of b^* to determine the decision function and establish three classification machine .At last we get the whole decision tree.

4.3 The Establishment of Classification Machine

1. Establish the C-SVC in heat fault and discharge fault.
2. Establish the C-SVC in low and high energy discharge fault.
3. Establish the C-SVC in low, middle and high temperature heating fault.

4.4 Test

4.4.1 Data test

Test the SVM model with the training sample and the test sample separately.

4.4.2 The Comparison between the Model and Three Ratio Methods

For the heating fault, three-ratio method is more precise. But it's less precise for discharge and low or middle heating fault. Overall, the three-ratio method there is still very inadequate in the gradual breakdown of fault. It is difficult for the diagnosis to continually deep and close with the real situation of the fault. But the SVM model can diagnose all types of fault and the fault segment with high reliability.

4.4.3 The Comparison between the Model and Neural Network

This paper uses the three-layer BP neural network, uses S-type function as hidden incentive function, uses a linear function as the output layer function and uses the

Table 3. The diagnosis results of the discharge fault and the heating fault accuracy rate

	time(s)	discharge (training)%	discharge (detected)%	heating (training)%	heating (detected)%
BP network	29.297	100	76	100	83
SVM model	2.828	96	83	95	87

	Accuracy rate (training) %	Accuracy rate (detected) %
BP network	100	79
SVM model	96	85

Table 4. The diagnosis results of the low and high energy discharge fault accuracy rate

	Time(s)	Low energy discharge fault (training) %	Low energy Discharge fault (detected)%	High energy discharge fault (training)%	High energy Discharge fault (detected)%
network	17.156	100	58	100	71
SVM model	2.219	90	83	87	88

	Accuracy rate (training) %	Accuracy rate (detected) %
BP network	100	66
SVM model	88	86

Table 5. Diagnosis results of the low middle and high temperature heating fault accuracy rate

	time(s)	Low middle temperature Heating fault (training) %	Low middle temperature Heating fault (detected) %	High temperature Heating fault (training) %	High temperature Heating fault (training) %
BP network	11.859	100	71	100	88
SVM model	1.563	88	100	93	88

	Accuracy rate (training) %	Accuracy rate (detected) %
BP network	100	66
SVM model	88	91

steepest descent gradient method to minimize error function. Test the BP network and the SVM model, and then we can get the following results with the same sample number.

Support Vector Machine is a learning method based on the structural risk minimization principle SRM, and it also minimizes the risk experience and model complexity. It ensures the model to have the best promotion ability and a smooth output function with limited samples. And the task of artificial neural network is to maintain the minimum risk experience. According to statistical Learning Theory, in a limited sample of cases, the experience risk minimization can not guarantee the actual risk minimization. This is the reason why there is an over-learning phenomenon in neural network. After testing the model which has been established, we can see that BP neural network model is only the whole sample study on the training sample, so the promotion of it is not high. On the contrary, support vector machine has a better applicability, and the model is closer to the real situation.

From the implementation time, the convergence speed of neural network is slow, and sometimes it is non- convergence. But support vector machines choose support vector only from the training samples and then establish support vector model directly according to support vector. This shows that the support vector machine has the advantage in the rapid diagnosis of transformer latent fault.

From the above discussion we can see that the support vector machines avoids the disadvantages of artificial neural network which are the low convergence speed slow, over-learning and local minimum. It is a more ideal transformer DGA fault analysis tools.

We can reduce the number of training samples while maintaining the same number of tested samples and the same parameters. Then, we can take a comparison between the two methods. Clearly, BP neural network method demands samples more, and the promotion of it is poor. But support vector machine method has the advantages of small sample study and the higher promotion.

5 Conclusions

The diagnosis accuracy rate of model to training samples and test samples is more than 80 percent. In addition, the judgment of model to training samples can not be 100% accurate judgment. The SVM rebuilds the high-dimensional space according to the distribution of support vector to finish the classification. It won't establish the classification model entirely according to all the training samples. It avoids the disadvantage of neural network "over-learning" effectively, and it has a strong outreach capacity. Through the comparison with the three-ratio method, it is proved that the SVM model is a more effective diagnostic tool in the deep diagnosis of fault. While through the comparison with the BP neural network method, the advantages about small sample study, stronger outreach capacity and an ability to avoid over-learning are fully reflected. From the implementation time, the convergence is faster, and the fault judgment time is shorter.

Acknowledgements. The project sponsored by the scientific research foundation for the returned overseas Chinese scholars, State Education Ministry, China.

References

1. Vapnik, V.N., Levin, E., Cun, Y.L.: Measuring the VC-Dimension of a Learning Machine. Neural Computation, 851–876 (1994)
2. Vapnik, V.N.: The Nature of Statistical Learning Theory, pp. 72–236. Springer, New York (1995)
3. Colin, M.: A Comparison of Photo - Acoustic Spectrometer and Gas Chromatograph Techniques for Dissolved Gas Analysis of Transformer Oil. In: ERPI Substation Equipment Diagnostics Conference, New Orleans, I. A, USA (2003)
4. Zhang, J., Walter, G.: Wavelet Neural Network for Function Learning. IEEE Trans on Signal Processing 43(6), 1485–1497 (1995)
5. Wang, Y.Z.: A combined ANN and Expert System Tool for Transformer Fault Diagnosis. IEEE Trans. On Power Delivery 13(4), 1148–1158 (1998)
6. Weston, J., Watkins, C.: Multi Class Support Vector Machines. Royal Holloway College. Tech Rep: CSD-TR-98-04 (1998)
7. Booth, R.R.: Power System Simulation Model Based on Probability Analysis. IEEE Transactions on Power Apparatus and Systems 1, 62–69 (1992)
8. Cardoso, A.J.M., Saraiva, E.S.: Computer Aided Detection of Air Gap Eccentricity in Operating Three Phase Induction Motors by Park's Vector Approach. IEEE Trans. on Industry Applications 29, 897–901 (1993)
9. Mejjari, H., Benbouzid, M.E.H.: Monitoring and Diagnosis of Induction Motors Electrical Faults using a Current Park's Vector Pattern Learning Approach. IEEE Trans. on Industry Applications 36, 730–735 (2000)
10. Cruz, S.M.A., Cardoso, A.J.M.: Rotor Cage Fault Diagnosis in Three Phase Induction Motors by Extended Park's Vector Approach. Electric Machines and Power Systems 28, 289–299 (2000)

A Test Data Compression Scheme for Reducing Power Based on OLELC and NBET

Jun Ma

Dept. of Educational Technology, Anqing Normal College

Abstract. A new test data compression scheme is introduced. This scheme encodes the test data provided by the core vendor, using a new and very effective compression scheme based on OLEL coding and neighboring bit-wise exclusive-or transform(OCNBET). Codeword is divided into two parts according to the position: odd bits and even bits. The odd bits of codeword are used to represent the length of runs and the even bits of codeword are used to represent whether a run is finished. Furthermore, a neighboring bit-wise exclusive-or transform is introduced to increase the probability of runs of 0s in the transformed data, so significant compression improvements and power efficient compared with the already known schemes are achieved. A simple architecture is proposed for decoding the compressed data on chip. Its hardware overhead is very low and comparable to that of the most efficient methods in the literature. Experimental results for the six largest ISCAS-89 benchmark circuits show that the proposed scheme is obviously better than the already known schemes in the aspects of compression ratio, power and the decompression structure.

Keywords: OLEL, Neighboring Bit-Wise Exclusive-or Transform(OCNBET), Data Compression.

1 Introduction

A System-on-a-Chip (SoC) chip, containing many modules and IP's on a single chip, has the advantage of reducing cost on fabrication, but suffers the disadvantage of increasing cost on testing due to increased complexity of test generation and large volume of test data. In addition, large amount of switching activity of scan test beyond that of the normal operation leads to power problem. Therefore, energy, peak power and average power during testing should be controlled to avoid causing chip malfunction and reducing reliability for circuit-under-test (CUT).

To cope with the problem of huge test data, many compression techniques and architectures had been proposed and developed. One type of these compression methods is to use codeword, for example, Golomb codes [1], Selective Huffman codes [2], VIHC codes [3], and FDR codes [4], etc to represent data block. A comprehensive study on these compression schemes was presented in [5] and the maximum compression of the methods of this type can be predicted by the entropy within test data. Another type of compression methods is to compress test data utilizing the bit correlation of test vectors to obtain minimum bit flips between consequent test patterns [6-8] to achieve test compression. An approach, OPMISR, uses ATE's repeat instructions and run-length coding to improve test compression [9].

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 728 – 735, 2008.
© Springer-Verlag Berlin Heidelberg 2008

As it was reported that, test patterns for large designs have very low percentage of specified bits, by exploiting that, high compression rate is achievable. For example, the hybrid test [10] approach generated both random and deterministic patterns for the tested chip while used an on-chip LFSR to generate the random patterns. The Broadcast (or Illinois) approach [11] used one pin to feed multiple scan chains. In [12, 13], a combinational network was used to compress test cubes to conceal large internal scan chains seen by the ATE. Also, several efficient methods such as RIN [14], SoCBIST [15], and EDT [16], etc, were proposed to achieve test data reduction by using an on-chip circuitry to expand compressed data to test patterns. Tang et al [17] also proposed a scheme to use switch configurations to deliver test patterns and provided a method to determine the optimum switch configuration.

2 Algorithm of OCNBET

2.1 OLEL Coding

In this section, the algorithm of OELE coding is described. It is a variable-to-variable length coding, which maps variable-length runs of 0s to variable-length codeword. The encoding procedure is shown as Table 1.

Table 1. The encoding procedure

Run length	Group	codes	Improved codes	codeword
0	A_1	10	0	01
1		11	1	11
2	A_2	100	00	00 01
3		101	01	00 11
4		110	10	10 01
5		111	11	10 11
6	A_3	1000	000	00 00 01
7		1001	001	00 00 11
……		……	……	……
12		1110	110	10 10 01
13		1111	111	10 10 11
……	……	……	……	……

The last column is the codeword of OELE coding. All of the odd bits of the codeword are corresponding improved codes. And all of the even bits of the codeword are labels which represent whether the codeword ends. When the label is 1, then it represents that the corresponding codeword ends. Or else, then it represents that the corresponding codeword continues. This characteristic is very useful for the decoding, which is introduced in section 3.

2.2 Don’t-Care Filling for Low-Power Scan Test

There are two possible types of test set: either every bit in a test vector is fully specified, or some bit are not specified. A test vector with unspecified bits is usually

referred to as a test cube. If the initial test set is not fully specified, it will be easier to compress. However, the size of recovered test data T_D will be larger, and thus the test application time is longer. An example of test set with test cubes is shown in Figure 1. Since the don't-care bits in test cubes can be randomly assigned to 0 or 1, it is possible to produce longer runs of 0s and 1s. As a result, the power during shift-in process can be greatly increased.

Test cubes	Test vectors
xxx00xxx1xxxxxxxx	00000000111111111
xxx10xxx1xxxxxxxx	11110000111111111
xxxx00100xxxxxx1x	00000010000000011

Fig. 1. Assign 0 or 1 to don't-care bits in test cubes

A greedy approach is adopted to assign don't-care bits from left to right, in which either '0' or '1' is selected to make the current run longer. According to this heuristic, the test cubes in are assigned to the test vectors as shown Figure 1.

2.3 Neighboring Bit-Wise Exclusive-or Transform（NBET）

In this section, we describe our transform technique which is the pre-processing method before compressing low-power scan test data to achieve high compression.

In our technique, first, don't-cares in a precomputed scan vectors are assigned by using heuristic to save scan test power, and then this yields the fully specified test set $T_D = \{V_1, V_2, V_3, ., V_m\}$, where V_i is the *i*th scan vector, having a lot of long runs of both 0s and 1s. Next, T_D is transformed to improve compression efficiency by using the neighboring bitwise exclusive-OR (NB-XOR) transform which performs XOR between neighboring bits on scan vector V_i $\{b_{i1}, b_{i2}, b_{i3},..., b_{i4}$, where b_{i1} is the *j*th bit in it, and this results in the NB-XORed difference vector set T_{NB} which is highly biased to 0s, and compression is carried out on it. T_{NB} is defined as follows:

$$T_{NB} = \{d_{NB1}, d_{NB2}, d_{NB3}, \cdots, d_{NBm}\}$$
$$d_{NBi} = \{b_{NBi1}, b_{NBi2}, b_{NBi3}, \cdots, b_{NBin-1}, b_{NBin}\}$$
$$= \{b_{i-10} \oplus b_{i1}, b_{i1} \oplus b_{i2}, b_{i2} \oplus b_{i3}, \cdots, b_{in-1} \oplus b_{in}\}$$

where d_{NBi} and b_{NB1j} are the *i*th NB-XORed difference vector and the *i*th NB-XORed difference bit in it, respectively. b_{NBij} is obtained by executing XOR between both neighboring bits $b_{i\,1}$ and b_{ij} in V_i.

Figure 2 shows an example of NB-XOR transform procedure sequenced by V, V_D and d_{NB}, where V, V_D and d_{NB} are a precomputed scan vector with don't-cares, a fully specified scan vector after don't-care assignment by MTC-filling and an NB-XORed difference vector we achieved at the end, respectively, and the left most bit for each vector is the first bit scanned into a scan chain with length 30. The total number of 1s in V_D is 21 and this is greatly reduced to a total of 3 as seen in d_{NB} by NB-XOR technique.

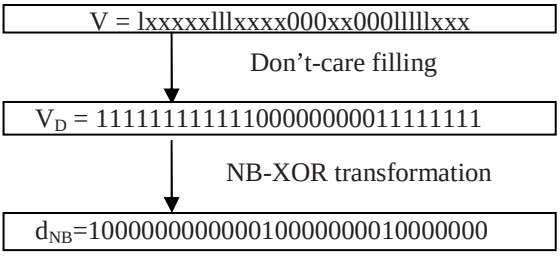


Fig. 2. Example of NB-XOR transform

3 Power Estimation

We use the weighted transitions metric(WTM) introduced in [4] to estimate the power consumption due to scan patterns. The WTM metric models the fact that scan power for a given vector depends not only on the number of transitions in it but also on their relative positions. WTM is also strongly correlated to the switching activity in the internal nodes of the circuit under test during scan operation. It was shown experimentally that scan vectors with higher weighted transition metric dissipate more power in the circuit under test. In addition, experimental results for an industrial ASIC confirm that the transition power in the scan chains is the dominant contributor to test power.

4 Design of Decoder

In this section, we illustrate the design of the decoder of the OCNBET.

From Table 1, if we add a prefix 1 to the odd bits(*b*) of the codeword, namely *1b*, then this value is 2 longer than real run-length(*l*). That is to say, *l=(1b)-2*. This is very useful. We can use this discipline to decompress.

This discipline can be implemented by a special counter. The special counter has two characteristics: (1) setting the lowest bit of the counter to number 1. When the data needing to be decoded is shifted into the counter, the number 1 and other data are together shifted to high bits of the counter. (2) using non-0 low bound. Traditional counter counts down to 0, but this counter counts down to 2 (the content of counter is "10"). This function can be implemented easily by some combinational circuits, which will not introduce high hardware overhead.

The design of this decoder is similar to that of the FDR using the FSM-based design. The decoder is embedded in chip, which has some characteristic of small size and simple structure. This decoder does not introduce hardware overhead obviously, and is independent of not only DUT, but precomputed test data as well. The block diagram of the decoder is shown in Figure 3. The decoder decodes the precomputed test data T_E to output T_D. The structure of this decoder is simple, which only needs a FSM and a special k+1 bits counter.

In Figure 3, to test the CUT with a single scan chain, first, the compressed NB-XORed difference vectors (T_E) from an ATE are transferred to an on-chip decoder

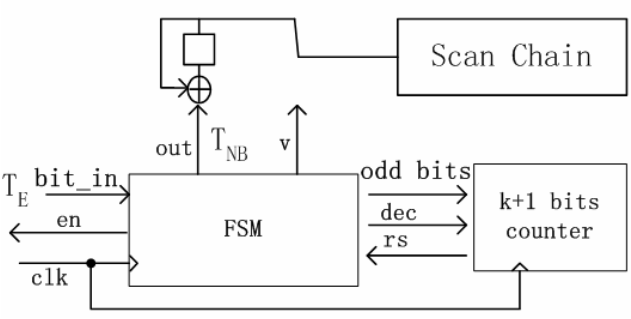


Fig. 3. Diagram of decoder structure

which decompresses them into the original NB-XORed difference vectors (T_{NB}). And then the decompressed vectors are fed into the inverse NB-XOR block, which consists of a simple pair of one XOR gate and one flipflop, to transform them into the original low-power scan vectors (T_D) under the assumption that the flipflop is initialized to 0. Finally, the transformed vectors are shifted into the internal scan chain. The serial output of this flip-flop feeds the serial input of the scan chain, and at the same time it loops back and is XORed with the serial output of the decoder.

The signal *bit_in* is primary input. The encoded test data is sent via *bit_in* to the FSM if the signal *en* is high. The signal *out* outputting decoded data is valid while the *v* is high. The *odd_bit* is a path which sends the odd bits of codeword to the k+1 bits counter. The *dec* is used to notify *k+1* bits counter to start counting down, and the *rs* is high as the counting finishes.

The basic principles of decoder working operation are as follows:

(1) At first, the FSM makes an enable signal named *en* into high level and *v* signal into low level, which indicates that the output signal *out* is invalid.
(2) Then FSM repeats to read the encoded data via *bit_in* until the value of even bits equal to 1 and send the odd bits of the encoded data to the *k+1* bits counter.
(3) After the special *k+1* bits counter receives the whole odd bits, the FSM sets *dec* and *v* high level and outputs several 0s till the k+1 bits counter finishes counting down. This time the value of k+1 bits counter equals to 2 (the content of this counter is “100”).
(4) Go to (2).

5 Experimental Results

In this section we will verify the effectiveness of the proposed scheme by using experiment results. For comparison with other schemes, we adopted the Mintest test sets [4] provided by Duke University of America. Most contributions on test data compression use the same circuits for experiments. Experiments were performed on the several largest ISCAS 89 benchmark circuits [18].

In order to compare the compression effect, the compression ratio is used. The compression ratio is computed as CR=(size of original data – size of compressed data)/ size of original data. The compression ratio of proposed scheme compared with Golomb is shown in Table 2.

Table 2. Compression obtained by using different codes methods

Circuit	Size of T_D bit	Proposed		Golomb[1]		MTC
		Size bit	CR %	m	CR %	
s5378	23754	**10621**	55.29	4	40.70	38.49
s9234	39273	**20085**	48.86	4	43.34	39.06
s13207	165200	**28543**	82.72	16	74.78	77.00
s15850	76986	**23857**	69.01	4	47.11	59.32
s38417	164736	**61847**	62.46	4	44.12	55.42
s38584	199104	**53098**	73.33	4	47.71	56.63
Avg.			65.28		49.63	54.32

In these experiments, we use runs of 0s. The 1st column of Table 2 is the circuit name. The 2nd column is the test size of original test data. The compression effect is shown in 3rd and 4th column of Table 2. The other columns are the compression ratio of Golomb and MTC. From Table 2, as can be seen, the proposed scheme achieves higher compression than any of the Golomb and MTC schemes for all the circuits. On average, the percentage compressions of the proposed scheme are 15.65% and 10.96% higher than those of Golomb and MTC.

Next we present the experimental results on the peak and average power consumption during the scan-in operation. These results show that test data compression can also lead to significant savings in power consumption. As described in Section 3, we estimate power using the weighted transitions metric. Let P^C_{peak} (P^C_{avg}) be the peak (average) power with compacted test sets obtained using Mintest. Similarly, let P^G_{peak} (P^G_{avg}) be the peak (average) power when Golomb coding is used by mapping the don't cares in T_D to zeros and let P^{SRLC}_{peak} (P^{SRLC}_{avg}) be the peak (average) power when SRLC coding is used by mapping the don't cares in T_D to zeros or ones according to our program.

Table 3 compares the average and peak power consumption for Mintest test sets when Golomb coding and our proposed scheme. The first column of Table 3 is the circuit name. The second and third columns are the peak power and average power if original Mintest test sets are used. The fourth and fifth columns are the peak power and average power if Golomb coding is used. The last two columns are peak power and average power if our proposed scheme is used. Table 3 shows that the peak power and average power are significantly less if proposed scheme is used for test data compression and the decompressed patterns are applied during testing. On

Table 3. Experimental results on peak and average scan-in power consumption

Circuit	Mintest		Golomb		Proposed	
	Peak power	Average Power	Peak power	Average Power	Peak power	Average Power
s9234	17494	14630	12994	5692	13720	4534
s13207	135607	122031	101127	12416	94886	8190
s15850	100228	90899	81832	20742	70908	13815
s35932	707280	583639	172834	73080	107226	40395
s38417	683765	601840	505295	172665	438005	118628
s38584	572618	535875	531321	136634	481185	86415
Avg.	369499	324819	234233	70205	200988	45329

average, the peak (power) is 33.41% (12.43%) less in our proposed scheme than for the Golomb coding.

In order to further demonstrate the performance of proposed scheme, the hardware overhead is analyzed. For group size A_i, the decoders of Golomb are formed from a FSM at least 9 states, and i bits counter; and the decoders of FDR are formed from a FSM, at least 9 states, $i+\log_2 i$ bits counter. Compared with the decoders of FDR and Golomb, the proposed scheme has only a FSM 2 states, k+1 bits counter (k is determined by the $2^k - 3 < l_{\max} \le 2^{k+1} - 3$). For all of these decoders on chip, with the different group size of 4, 8, 16, the statistical and comparative data of the hardware overhead are illustrated in Table 4. All of these circuits are synthesized with the synopsys design compiler and mapped to the 2-input nand cell of the TSMC35 library. As shown in Table 5, the hardware overhead of Golomb and FDR are larger than that of OLEL coding.

Table 4. Hardware overhead comparison

Compression method	Area overhead (map to 2-input nand gate)			
Group size	2	4	8	16
Golomb	74	125	227	307
FDR			320	
OLEL coding			66	

6 Conclusion

A new test compression/decompression technique is introduced to reduce test data volume and scan test power consumption, simultaneously. In the experimental results, considerable reduction of test data volume and scan test power is achieved. Therefore, the proposed approach can provided significant reductions in test data volume and, at the same tem, the power consumption.

References

1. Chandra, C., Chakrabarty, K.: System-On-A-Chip Test-Data Compression and Decompression Architectures Based on Golomb Codes. IEEE CAD/ICAS 20, 355–368 (2001)
2. Jas, J., Ghosh, D., Touba, N.: An Efficient Test Vector Compression Scheme Using Selective Huffman Coding. IEEE CAD 21, 797–806 (2003)
3. Gonciari, B., Hashimi, A., Nicolici, G.: Variable-Length Input Huffman Coding for Systemon-a-chip Test. IEEE CAD 22, 783–796 (2003)
4. Chandra, C., Chakrabarty, K.: Frequency-Directed Run-Length Codes with Application to System-on-a-chip Test Data Compression. In: Proceedings of VTS, pp. 42–47 (2001)
5. Balakrishnan, B., Touba, N.: Relating Entropy Theory to Test Data Compression. In: Proceedings of ETS, pp. 94–99 (2004)
6. Reda, C., Orailoglu, J.: Reducing Test Application Time through Test Data Mutation Encoding. In: Proceeding of DATE, pp. 387–393 (2002)
7. Jas, J., Touba, N.: Using an Embedded Processor for Efficient Deterministic Testing of Systems-on-a-Chip. In: Proceedings of ICCD, pp. 418–423 (1999)
8. Lin, C., Chen, L.: A Cocktail Approach On Random Access Scan Toward Low Power and High Efficiency Test. In: Proceedings of ICCAD, pp. 94–99 (2005)
9. Barnhart, D., Distler, O., Farnsworth, A., Ferko, B.: Extending OPMISR beyond 10x Scan Test Efficiency. IEEE Design & Test Compututer, 1965–1973 (2002)
10. Jas, J., Touba, N.: Test Data Compression Technique for Embedded Cores Using Virtual Scan Chains. IEEE VLSI 12, 775–780 (2004)
11. Pandey, A., Patel, J.: An Incremental Algorithm For Test Generation in Illinois Scan Architecture Based Designs. In: Proceedings of DATE, pp. 368–375 (2002)
12. Bayraktaroglu, I., Orailoglu, A.: Test Volume and Application Time Reduction Through Scan Chain Concealment. In: Proceedings of DAC, pp. 151–155 (2001)
13. Bayraktaroglu, I., Orailoglu, A.: Decompression Hardware Determination for Test Volume and Time Reduction Through Unified Test Pattern Compaction and Compression. In: Proceedings of VTS, pp. 113–120 (2003)
14. Li, L., Chakrabarty, K.: Deterministic BIST Based on a Reconfigurable Interconnection Network. In: Proceedings of ITC, pp. 460–469 (2003)
15. Synopsys Inc., http://www.synopsys.com/
16. Rajski, J.: Embedded Deterministic Test. IEEE TCAD 23, 776–792 (2004)
17. Tang, H., Reddy, S., Pomeranz, I.: On Reducing Test Data Volume and Test Application Time for Multiple Scan Chain Designs. In: Proceedings of ITC, pp. 1079–1088 (2003)
18. Brglez, F.: Combinational Profiles of Sequential Benchmark Circuits. In: Proc of 1989 International Symposium on Circuits and Systems, pp. 1929–1934 (1989)

On Initial Rectifying Learning for Linear Time-Invariant Systems with Rank-Defective Markov Parameters

Mingxuan Sun

College of Information Engineering
Zhejiang University of Technology
Hangzhou 310014, China
mxsun@zjut.edu.cn

Abstract. This paper presents an initial rectifying learning method for trajectory tracking of linear time-invariant systems with rank-defective Markov parameters. The initial shift problem is addressed through introduction of the initial rectifying action. The role of the rectifying action is examined in case of systems with row and column rank-defective Markov parameters, respectively. Sufficient conditions for convergence of the proposed learning algorithms are derived, by which the learning gains can be chosen. It is shown that the output trajectory converges to the desired one with a smooth transition. The merging of the transition to the desired trajectory occurs at a pre-specified time instant.

Keywords: Initial shift problem, convergence, learning algorithms, rank-defective Markov parameters.

1 Introduction

Iterative learning is a methodology suitable for a system performing identical tasks of a finite duration repetitively, in the manner analogous to industrial manipulators operations [1]. At the beginning of each cycle, the system is repositioned to a *home* position, and the input is applied, which is determined on the basis of data from the last cycle. Learning algorithms need less *a priori* knowledge about the system dynamics and involve much less on-line computational effort. There have been efforts made to clarify the principle of iterative learning methodology. In [2], it was showed that the direct transmission term of the input to the output plays a crucial role in the convergence of learning processes. For systems with well-defined relative degree, it was delineated that the highest order of the error derivative used in the learning algorithm should be equal to the relative degree. This fundamental property has been examined further in [3]. In addition, in [4], iterative learning was characterized for systems with rank-defective Markov parameters.

The advantages offered by iterative learning lie in the resultant transient behavior. It is a common requirement that the initial condition at each cycle is reset

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 736–743, 2008.
© Springer-Verlag Berlin Heidelberg 2008

to the initial condition corresponding to the desired trajectory. Iterative learning mainly benefits from this requirement to realize complete tracking along the entire span of operation interval [5]. In the published literature, a milder requirement is made for the case where an initial shift exists, i.e., the initial condition at each cycle remains the same but different from the desired initial condition [6]. The developed learning algorithm enables the system to possess asymptotic tracking capability. Intuitively, the impulsive action is required in learning algorithms to totally eliminate the effect due to an initial shift. In [7], such a learning algorithm was given, and it is verified to be able to achieve complete tracking. However, the use of the impulsive action is not practical. The initial rectifying action was suggested in [8,9,10], and was shown to be effective to guarantee convergence of the learning process to a reasonable outcome. The system output is ensured to converge to the desired one on a pre-specified interval with a predetermined transient response. Compared with the initial impulsive action [7], the initial rectifying action is finite and implementable.

This paper addresses the initial shift problem for irregular linear time-invariant systems, i.e. systems with rank-defective 0th, 1st, $\cdots$, and $(p-1)$th Markov parameters but full-rank pth Markov parameters ($p = 1, 2, \cdots$), which represents a continuation of previous works [7,8]. In this paper, more general classes of systems are considered by distinguishing row or column rank-defective cases. Theoretical results are presented to demonstrate the learning performance of the proposed initial rectifying action in dealing with initial shift.

2 Problem Formulation and Preliminaries

Consider the class of linear time-invariant systems described by

$$\dot{x}(t) = Ax(t) + Bu(t) \,, \tag{1}$$

$$y(t) = Cx(t) + Du(t) \,. \tag{2}$$

where $A \in R^{n\times n}, B \in R^{n\times r}, C \in R^{m\times n}$, and $D \in R^{m\times r}$ are constant matrices, $x(t) \in R^n$, $u(t) \in R^r$, and $y(t) \in R^m$ are, respectively, the state, control input, and output of the system. The objective of this paper is to find an input profile such that the system output follows the given desired trajectory as closely as possible. We consider the case where the system is not reset to the desired starting point at each cycle. Instead, there exists an initial shift between the resetting position and the desired initial position, i.e., $\|x_d(0) - x_k(0)\| \geq c_{x_{d0}}$, where $c_{x_{d0}}$ is a small positive real number. One example is that the system output is expected to track a step command from the resetting zero position. We shall introduce an initial rectifying action in the learning algorithm to overcome such a substantial large initial shift. The initial rectifying action will produce a smooth transience from the system resetting point to the desired trajectory. The merging time is a design parameter specified by the initial rectifying action.

In this paper, this objective is achieved for the systems with rank-defective Markov parameters. To this end, let us define two classes of systems we wish to consider.

i) The system (1)-(2) has row rank-defective Markov parameters:

$$\begin{aligned} 0 \le \operatorname{rank}(D) &< m \\ 0 \le \operatorname{rank}(CA^{i-1}B) &< m, i = 1, 2, \cdots, p-1 \\ \operatorname{rank}(CA^{p-1}B) &= m \ , \end{aligned} \tag{3}$$

where $p \ge 1$; and
ii) The system (1)-(2) has column rank-defective Markov parameters:

$$\begin{aligned} 0 \le \operatorname{rank}(D) &< r \\ 0 \le \operatorname{rank}(CA^{i-1}B) &< r, i = 1, 2, \cdots, p-1 \\ \operatorname{rank}(CA^{p-1}B) &= r \ , \end{aligned} \tag{4}$$

where $p \ge 1$.

To design the initial rectifying action, we need the definition for the function $\theta_{p,h} : [0, T] \to R$, which is defined as

$$\theta_{p,h}(t) = \begin{cases} \frac{1}{h^{2p+1}} \frac{(2p+1)!}{p!^2} t^p (h-t)^p, & t \in [0, h] \\ 0, & t \in (h, T] \end{cases} \tag{5}$$

and satisfies

$$\int_0^h \theta_{p,h}(s) ds = 1 \tag{6}$$

$$\left(\frac{t^j}{j!} \int_t^h \theta_{p,h}(\tau) d\tau \right)^{(q)} \Bigg|_{t=0} = \begin{cases} 1, q = j \\ 0, q \ne j \end{cases} \tag{7}$$

where h is a design parameter.

3 Systems with Row Rank-Defective Markov Parameters

For the system (1)-(2) with row rank-defective Markov parameters, an updating law with initial rectifying action is proposed in the form of

$$\begin{aligned} u_{k+1}(t) = u_k(t) &+ \sum_{j=0}^{p} \Gamma_j \left(y_d^{(j)}(t) - y_k^{(j)}(t) \right) \\ &- \sum_{j=0}^{p-1} \sum_{q=0}^{p} \left(\frac{t^j}{j!} \int_t^h \theta_{p,h}(s) ds \right)^{(q)} \Gamma_q \left(y_d^{(j)}(0) - y_k^{(j)}(0) \right) \ , \end{aligned} \tag{8}$$

where $t \in [0, T]$; k refers to the number of operation cycle; $y_d(t), t \in [0, T]$ is the desired trajectory; and $\Gamma_j, j = 0, 1, \cdots, p$, are the gain matrices.

Let $\rho(H)$ denote the spectral radius of the matrix H and $\lambda-$norm of a vector-valued function $h(t)$ be defined as $\|h(\cdot)\|_\lambda = \sup_{t \in [0,T]} e^{-\lambda t} \|h(t)\|$, where $\| \cdot \|$ is an appropriate vector norm. The following theorem states the convergence result when applying the updating law (8).

Theorem 1. Given a desired trajectory $y_d(t), t \in [0,T]$, which is p times continuously differentiable, let the updating law (8) be applied to the system (1)-(2) with row rank-defective Markov parameters, defined as (3). If

$$\text{(C1) } D\Gamma_i + \sum_{j=i+1}^{p} CA^{j-i-1}B\Gamma_j = 0, i = 1, 2, \cdots, p$$

$$\text{(C2) } \rho\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right) < 1$$

$$\text{(C3) } x_k(0) = x^0$$

then $y_k(t) \to y^*(t)$ uniformly on $[0,T]$ as $k \to \infty$, where

$$y^*(t) = y_d(t) - \sum_{j=0}^{p-1}\left(\frac{t^j}{j!}\int_t^h \theta_{p,h}(s)ds\right)\left(y_d^{(j)}(0) - y_0^{(j)}(0)\right) \tag{9}$$

Proof. By the property of $\theta_{p,h}(t)$ described by (7), it follows that $y^{*(j)}(0) = y_0^{(j)}(0), j = 0, 1, \cdots, p-1$, which results in $u_{k+1}(t) = u_k(t) + \sum_{j=0}^{p}\Gamma_j e_k^{*(j)}(t) - \sum_{j=0}^{p-1}\sum_{q=0}^{p}\left(\frac{t^j}{j!}\int_t^h \theta_{p,h}(s)ds\right)^{(q)}\Gamma_q e_k^{*(j)}(0)$ where $e_k^*(t) = y^*(t) - y_k(t)$. Let $u^*(t)$ be a control input such that $y^*(t) = Ce^{At}x^0 + C\int_0^t e^{A(t-s)}Bu^*(s)ds + Du^*(t)$. Then, using (C1), the error $e_k^*(t)$ for the $(k+1)$th cycle can be written as

$$\begin{aligned}
e_{k+1}^*(t) = &\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right)e_k^*(t) - \sum_{j=0}^{p}\int_0^t Ce^{A(t-s)}A^jB\Gamma_j e_k^*(s)ds \\
&+ \sum_{j=1}^{p}\sum_{q=j}^{p} Ce^{At}A^{q-j}B\Gamma_q e_k^{*(j-1)}(0) \\
&+ \sum_{j=0}^{p-1}\sum_{q=0}^{p}\int_0^t Ce^{A(t-s)}B\left(\frac{s^j}{j!}\int_s^h \theta_{p,h}(\tau)d\tau\right)^{(q)}\Gamma_q e_k^{*(j)}(0)ds \\
&+ \sum_{j=0}^{p-1}\sum_{q=0}^{p}\left(\frac{t^j}{j!}\int_t^h \theta_{p,h}(s)ds\right)^{(q)} D\Gamma_q e_k^{*(j)}(0)\,.
\end{aligned} \tag{10}$$

It is seen that $e_0^{*(j)}(0) = 0, j = 0, 1, \cdots, p-1$, and thus $e_k^{*(j)}(0) = 0, j = 0, 1, \cdots, p-1$, by which (10) becomes

$$\begin{aligned}
e_{k+1}^*(t) = &\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right)e_k^*(t) \\
&- \sum_{j=0}^{p}\int_0^t Ce^{A(t-s)}A^jB\Gamma_j e_k^*(s)ds\,.
\end{aligned} \tag{11}$$

As is known [11], for any given $\varepsilon > 0$ and an appropriate norm $\|\cdot\|$, there exists an invertible matrix $Q \in R^{m\times m}$ such that $\|Q\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right)Q^{-1}\| \leq \rho\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right) + \varepsilon$. Thus, taking norms of both sides of (11) yields

$$\begin{aligned}\|Qe_{k+1}^*(t)\| \leq & \left[\rho\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right) + \varepsilon\right]\|Qe_k^*(t)\| \\ & + \int_0^t \|\sum_{j=0}^{p} QCe^{A(t-s)}A^jB\Gamma_jQ^{-1}\|\|Qe_k^*(s)\|ds\end{aligned}$$

Multiplying both sides by $e^{-\lambda t}$ and using the definition of λ-norm, we have

$$\begin{aligned}&\|Qe_{k+1}^*(\cdot)\|_\lambda \\ &\leq \left[\rho\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right) + \varepsilon + b\frac{1-e^{-\lambda T}}{\lambda}\right]\|Qe_k^*(\cdot)\|_\lambda\,, \quad (12)\end{aligned}$$

where $b = \sup_{t\in[0,T]}\|\sum_{j=0}^{p} QCe^{At}A^jB\Gamma_jQ^{-1}\|$. By using (C2) and that $\varepsilon > 0$ can be a value small enough, it is possible to find a λ sufficiently large such that $\rho\left(I - D\Gamma_0 - \sum_{j=1}^{p} CA^{j-1}B\Gamma_j\right) + \varepsilon + b\frac{1-e^{-\lambda T}}{\lambda} < 1$. Therefore, (12) is a contraction in $\|Qe_k^*(\cdot)\|_\lambda$ and $\|Qe_k^*(\cdot)\|_\lambda \to 0$, as $k \to \infty$. By the definition of $\lambda-$norm and the invertibility of the matrix Q, we obtain $\sup_{t\in[0,T]}\|e_k^*(t)\| \leq e^{\lambda T}\|Q^{-1}\|\|Qe_k^*(\cdot)\|_\lambda$. Then, $y_k(t) \to y^*(t)$ uniformly on $[0,T]$ as $k \to \infty$. □

Theorem 1 implies that a suitable choice for the gain matrices leads to the convergence of the system output to the trajectory $y^*(t)$ for all $t \in [0,T]$, as the repetition increases. To refer to the definition of $y^*(t)$, it is derived that $y^*(t) = y_d(t), t \in (h,T]$. Therefore, the uniform convergence of the system output to the desired one is achieved over the interval (h,T], while the output trajectory on $[0,h]$ is produced by the initial rectifying action, and it is a smooth transition from the initial position to the desired trajectory. The initial transient trajectory joins the desired trajectory at time moment $t = h$, which can be pre-specified by the designer.

4 Systems with Column Rank-Defective Markov Parameters

In Section 3, the systems with row rank-defective Markov parameters are considered, which requires that the number of outputs is not smaller than the number of inputs. Now, we consider the systems with column rank-defective Markov parameters. For this case, it allows that the number of outputs is less or equal to the number of inputs.

Replacing the term $y_d^{(j)}(0) - y_k^{(j)}(0)$ in (8) with $y_d^{(j)}(0) - y_0^{(j)}(0)$, a modified updating law is obtained as follows:

$$u_{k+1}(t) = u_k(t) + \sum_{j=0}^{p} \Gamma_j \left(y_d^{(j)}(t) - y_k^{(j)}(t) \right) - \sum_{j=0}^{p-1} \sum_{q=0}^{p} \left(\frac{t^j}{j!} \int_t^h \theta_{p,h}(s) ds \right)^{(q)} \Gamma_q \left(y_d^{(j)}(0) - y_0^{(j)}(0) \right) . \quad (13)$$

The sufficient condition for convergence of learning algorithm (13) is given in the following theorem.

Theorem 2. Given a desired trajectory $y_d(t), t \in [0, T]$, which is p times continuously differentiable, let the updating law (13) be applied to the system (1)-(2) with column rank-defective Markov parameters, defined as (4). If

$$\text{(C4) } \Gamma_i D + \sum_{j=i+1}^{p} \Gamma_j C A^{j-i-1} B = 0, i = 1, 2, \cdots, p$$

$$\text{(C5) } \rho \left(I - \Gamma_0 D - \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) < 1$$

$$\text{(C6) } x_k(0) = x^0$$

then $y_k(t) \to y^*(t)$ uniformly on $[0, T]$ as $k \to \infty$, where

$$y^*(t) = y_d(t) - \sum_{j=0}^{p-1} \left(\frac{t^j}{j!} \int_t^h \theta_{p,h}(s) ds \right) \left(y_d^{(j)}(0) - y_0^{(j)}(0) \right) \quad (14)$$

Proof. It follows from (7) that $y^{*(j)}(0) = y_0^{(j)}(0), j = 0, 1, \cdots, p-1$, which results in $u_{k+1}(t) = u_k(t) + \sum_{j=0}^{p} \Gamma_j e_k^{*(j)}(t)$, where $e_k^*(t) = y^*(t) - y_k(t)$. Let $u^*(t)$ be a control input such that $y^*(t) = Ce^{At}x^0 + C \int_0^t e^{A(t-s)} B u_d^*(s) ds + D u^*(t)$, and denote by $\Delta u_k^*(t) = u^*(t) - u_k(t)$ the input error. By the condition (C6), we have

$$e_k^*(t) = C \int_0^t e^{A(t-s)} B \Delta u_k^*(s) ds + D \Delta u_k^*(t) . \quad (15)$$

Hence, using the condition (C4) gives rise to

$$\sum_{j=0}^{p} \Gamma_j e_k^{*(j)}(t) = \left(\Gamma_0 D + \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) \Delta u_k^*(t) + \sum_{j=0}^{p} \Gamma_j C A^j \int_0^t e^{A(t-s)} B \Delta u_k^*(s) ds , \quad (16)$$

and the input error for the $(k+1)$th cycle can be written as

$$\Delta u_{k+1}^*(t) = \left(I - \Gamma_0 D - \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) \Delta u_k^*(t) - \sum_{j=0}^{p} \Gamma_j C A^j \int_0^t e^{A(t-s)} B \Delta u_k^*(s) ds \, . \tag{17}$$

For any given $\varepsilon > 0$ and an appropriate norm $\| \cdot \|$, there exists an invertible matrix $Q \in R^{m \times m}$ such that $\|Q \left(I - \Gamma_0 D - \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) Q^{-1}\| \leq \rho \left(I - \Gamma_0 D - \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) + \varepsilon$. Thus, taking norms of both sides of (17) yields

$$\|Q \Delta u_{k+1}^*(t)\| \leq \left[\rho \left(I - \Gamma_0 D - \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) + \varepsilon \right] \|Q \Delta u_k^*(t)\| + \int_0^t \| \sum_{j=0}^{p} Q \Gamma_j C A^j e^{A(t-s)} B Q^{-1} \| \|Q \Delta u_k^*(s)\| ds$$

Multiplying both sides by $e^{-\lambda t}$ and using the definition of λ-norm, we have

$$\|Q \Delta u_{k+1}^*(\cdot)\|_\lambda \leq \left[\rho \left(I - \Gamma_0 D - \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) + \varepsilon + b \frac{1 - e^{-\lambda T}}{\lambda} \right] \|Q \Delta u_k^*(\cdot)\|_\lambda \, , \tag{18}$$

where $b = \sup_{t \in [0,T]} \| \sum_{j=0}^{p} Q \Gamma_j C A^j e^{At} B Q^{-1} \|$. By (C5) and that $\varepsilon > 0$ can be a value small enough, it is possible to find a λ sufficiently large such that $\rho \left(I - \Gamma_0 D - \sum_{j=1}^{p} \Gamma_j C A^{j-1} B \right) + \varepsilon + b \frac{1 - e^{-\lambda T}}{\lambda} < 1$. Therefore, (18) indicates a contraction in $\|Q \Delta u_k(\cdot)\|_\lambda$, and $\|Q \Delta u_k^*(\cdot)\|_\lambda \to 0$, as $k \to \infty$. By the definition of $\lambda-$norm and the invertibility of the matrix Q, we obtain $\sup_{t \in [0,T]} \|\Delta u_k^*(t)\| \leq e^{\lambda T} \|Q^{-1}\| \|Q \Delta u_k^*(\cdot)\|_\lambda$. Then, $u_k(t) \to u^*(t)$ on $[0,T]$ as $k \to \infty$. Further, multiplying both sides of (15) by $e^{-\lambda t}$ gives $\|e_k^*(\cdot)\|_\lambda \leq \left(\frac{1 - e^{-\lambda T}}{\lambda} c + d \right) \|Q \Delta u_k^*(\cdot)\|_\lambda$, where $c = \sup_{t \in [0,T]} \|C e^{At} B Q^{-1}\|$ and $d = \|D Q^{-1}\|$. It is evident that $y_k(t) \to y^*(t)$ on $[0,T]$ as $k \to \infty$. □

In [7], the Dirac delta function is adopted and the proposed algorithm involves an impulsive action at the starting moment. As is well known, the ideal impulsive action cannot be implemented in practice. The initial rectifying action, however, is finite, which makes the learning algorithm practical. In addition, the initial positioning in [7] was required to satisfy $y_0(0) = y_d(0)$ and $x_{k+1}(0) = x_k(0)$. From Theorems 1 and 2, it is seen that the identical repositioning requirement is removed.

5 Conclusion

In this paper, in the presence of an initial shift, the iterative learning method has been presented for trajectory tracking of linear time-invariant systems with rank-defective Markov parameters. The initial rectifying learning algorithms have been proposed to overcome the effect of the initial shift, and sufficient conditions for convergence of the proposed learning algorithms have been derived. It has been shown that the system output is ensured to converge to the desired trajectory uniformly, jointed smoothly with a piece of transient trajectory from the starting position.

Acknowledgements. This research is supported by the National Natural Science Foundation of China (No. 60474005), and the Natural Science Foundation of Zhejiang Province (No. Y107494).

References

1. Arimoto, S.: Learning control theory for robotic motion. International Journal of Adaptive control and Signal Processing 4, 543–564 (1990)
2. Sugie, T., Ono, T.: An iterative learning control law for dynamical systems. Automatica 27, 729–732 (1991)
3. Ahn, H.-S., Choi, C.-H., Kim, K.-B.: Iterative learning control for a class of nonlinear systems. Automatica 29, 1575–1578 (1993)
4. Porter, B., Mohamed, S.S.: Iterative learning control of partially irregular multivariable plants with initial state shifting. International Journal of Systems Science 22, 229–235 (1991)
5. Heinzinger, G., Fenwick, D., Paden, B., Miyazaki, F.: Robust learning control. In: Proceedings of the 28th IEEE Conference on Decision and Control, Tempa FL, pp. 436–440 (1989)
6. Lee, H.-S., Bien, Z.: Study on robustness of iterative learning control with non-zero initial error. International Journal of Control 64, 345–359 (1996)
7. Porter, B., Mohamed, S.S.: Iterative learning control of partially irregular multivariable plants with initial impulsive action. International Journal of Systems Science 22, 447–454 (1991)
8. Sun, M., Huang, B.: Iterative Learning Control. National Defence Industrial Press, Beijing (1999)
9. Sun, M., Wang, D.: Initial condition issues on iterative learning control for nonlinear systems with time delay. International Journal of Systems Science 32, 1365–1375 (2001)
10. Sun, M., Wang, D.: Iterative learning control with initial rectifying action. Automatica 38, 1177–1182 (2002)
11. Ortega, J.M., Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables. Academic Press, New York (1970)

A New Mechanical Algorithm for Solving System of Fredholm Integral Equation Using Resolvent Method

Weiming Wang[1,⋆], Yezhi Lin[1], and Zhenbing Zeng[2]

[1] Institute of Nonlinear Analysis, College of Mathematics and Information Science, Wenzhou University, Wenzhou, 325035
[2] Software Engineering Institute of East China Normal University Shanghai 200062
weimingwang2003@163.com, lyz_eyes@163.com, zbzeng@sei.ecnu.edu.cn

Abstract. In this paper, by using the theories and methods of mathematical analysis and computer algebra, a complete iterated collocation method for solving system of Fredholm integral equation is established, furthermore, the truncated error is discussed. And a new mechanical algorithm **FredEqns** is established, too. The algorithm can provide the approximate solution of system of Fredholm integral equation. Some examples are presented to illustrate the efficiency and accuracy of the algorithm. This will be useful for the integral system solving and mathematical theorems automatic proving.

Keywords: Fredholm integral equation, Mechanical algorithm, Maple.

1 Introduction

It is currently very much in vogue to study the integral equation. Integral equation is an important branch of modern mathematics. Such equation occur in various areas of applied mathematics, physics, and engineering, etc [1,2,3].

Fredholm integral equation is one of important types of integral equation, which arises in various branches of applications such as solid mechanics, phase transitions, potential theory and Dirichlet problems, electrostatics, mathematical problems of radiative equilibrium, the particle transport problems of astrophysics and reactor theory, and radiative heat transfer problems, population dynamics, continuum mechanics of materials with theory, economic problems, non-local problems of diffusion and heat conduct and many others [1]. Fredholm integral equation has been extensively investigated theoretically and numerically in recent year.

It's well known that a computational approach to the solution of integral equation is, therefore, an essential branch of scientific inquiry [1]. And a computational approach to the solution of integral equation is an essential branch of scientific inquiry. As a matter of fact, some valid methods of solving Fredholm integral equation have been developed in recent years, such as quadrature

⋆ Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 744–754, 2008.
© Springer-Verlag Berlin Heidelberg 2008

method, collocation method and Galerkin method, expansion method, product-integration method, deferred correction method, graded mesh method, Sinc-collocation method, Trefftz method, Taylor's series method, Tau method, interpolation and Gauss quadrature rules and Petrov-Galerkin method, decomposition method, etc [1,2,3,4,5,6,7,8,9,10,11,12,13]. Besides, the resolvent method [1,2,3] is one of commonly used methods for solving the second kind of Fredholm integral equation.

However, for solving system of Fredholm integral equation, there are few literatures. In [1,2,3], P.K.Kythe, Y.Shen, S.Zhang had mentioned the problem of solving system of Fredholm integral equation respectively, involved the resolvent method. But they neither given the details of the method, nor illustrated any example. The objective of this paper is to establish the algorithm of iterated collocation method in details and establish a new mechanical algorithm in `Maple` for solving system of Fredholm integral equation.

We organize this paper as follows. In the next section, we recall the resolvent method for solving the system of Fredholm integral equation. In section 3, we establish a new algorithm for solving the system of Fredholm integral equation with mechanization. In section 4, we present some examples to illustrate the mechanical process and finally in section 5, we will show conclusion and remarks of this study.

2 Basic Methods

Consider the system of Fredholm integral equation as the form

$$x_i(s) - \lambda \sum_{j=1}^{n} \int_a^b k_{ij}(s,t)x_j(t)dt = y_i(s), \quad i = 1, 2, \cdots, n. \tag{1}$$

where $k_{ij}(s,t)$ $(i,j = 1, 2, \cdots, n)$ are the L_2-kernels of the equations (1), $y_i(s) \in L_2[a,b]$ are given, which satisfy

$$\int_a^b \int_a^b |k_{ij}(s,t)|^2 dsdt < \infty, \quad \int_a^b |y_j(t)|^2 dt < \infty,$$

λ is a parameter and satisfies

$$|\lambda| < \left\{ \max_{1\le i\le n} \sqrt{\int_a^b \int_a^b |k_{ij}(s,t)|^2 dsdt} \right\}^{-1},$$

and $x_i(s)$ $(i = 1, 2, \cdots, n)$ are the solutions to be determined.

Rewrite the system (1) in the form

$$X(s) - \lambda \int_a^{nb-(n-1)a} K(s,t)X(t)dt = Y(s), \tag{2}$$

where

$$X(s) = (x_1(s), x_2(s-(b-a)), \cdots, x_n(s-(n-1)(b-a)))^T,$$
$$Y(s) = (y_1(s), y_2(s-(b-a)), \cdots, y_n(s-(n-1)(b-a)))^T,$$

$$K(s,t) = \begin{pmatrix} k_{11}(s,t) & \cdots & k_{1n}(s,t-(n-1)(b-a)) \\ k_{21}(s-(b-a),t) & \cdots & k_{2n}(s-(b-a),t-(n-1)(b-a)) \\ \cdots & \cdots & \cdots \\ k_{n1}(s-(n-1)(b-a),t) & \cdots & k_{nn}(s-(n-1)(b-a),t-(n-1)(b-a)) \end{pmatrix}.$$

Then we can obtain the following theorem.

Theorem 1: If $y_i(s) \in [a,b]$ and $k_{ij}(s,t) \in L_2[a,b] \times [a,b]$ $(i,j = 1,2,\cdots,n)$, then $Y(s) \in L_2[a, nb-(n-1)a]$ and the kernel $K(s,t) \in L_2[a, nb-(n-1)a] \times [a, nb-(n-1)a]$.

Proof:

$$\begin{aligned} &\int_a^{nb-(n-1)a} \int_a^{nb-(n-1)a} |K(s,t)|^2 ds dt \\ &= \sum_{i,k=1}^{n} \int_{(m-1)b-(m-2)a}^{mb-(m-1)a} \int_{(i-1)b-(i-2)a}^{ib-(i-1)a} |K(s,t)|^2 ds dt \\ &= \sum_{i,m=1}^{n} \int_a^b \int_a^b |K(s+(m-1)(b-1), t+(i-1)(b-a))|^2 ds dt \\ &= \sum_{i,m=1}^{n} \int_a^b \int_a^b |k_{im}(s,t)|^2 ds dt < \infty. \end{aligned}$$

By the same taken, we can get

$$\int_a^{nb-(n-1)a} |Y(s)|^2 ds = \sum_{i=1}^{n} \int_a^b |y_i(s)|^2 ds < \infty.$$

Following, we give the basic resolvent method for solving system (2).

Theorem 2: Let $Y(s) \in L_2[a,b]$ and the kernel $K(s,t) \in L_2[a, nb-(n-1)a] \times [a, nb-(n-1)a]$ be continuous functions. Then the solution of (2) is given by

$$X(s) = Y(s) + \lambda \int_a^{nb-(n-1)a} K_\lambda(s,t) X(t)\, dt, \tag{3}$$

where

$$K_\lambda(s,t) = \frac{D_\lambda(s,t)}{d(\lambda)}, \tag{4}$$

the Fredholm determinant $d(\lambda)$ of the kernel $K(s,t)$ is:

$$d(\lambda) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \lambda^n B_n, \tag{5}$$

and $D_\lambda(s,t)$, known as the Fredholm minor, is as the form:

$$D_\lambda(s,t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \lambda^n A_n(s,t), \tag{6}$$

such that

$$
\begin{aligned}
&A_0 = k(s,t),\\
&A_n(s,t) = \underbrace{\int_a^{nb-(n-1)a} \cdots \int_a^{nb-(n-1)a}}_{n\ times} \begin{vmatrix} K(s,t) & \cdots & K(s,s_n) \\ K(s_1,t) & \cdots & K(s_1,s_n) \\ \cdots & \cdots & \cdots \\ K(s_n,t) & \cdots & K(s_n,s_n) \end{vmatrix} ds_1 \cdots ds_n,\\
&B_0 = 1,\\
&B_n = \underbrace{\int_a^{nb-(n-1)a} \cdots \int_a^{nb-(n-1)a}}_{n\ times} \begin{vmatrix} K(s_1,s_1) & \cdots & K(s_1,s_n) \\ K(s_2,s_1) & \cdots & K(s_2,s_n) \\ \cdots & \cdots & \cdots \\ K(s_n,s_1) & \cdots & K(s_n,s_n) \end{vmatrix} ds_1 \cdots ds_n\\
&(n = 1, 2, \cdots)
\end{aligned}
\tag{7}
$$

In most problems the formula (7) is not very convenient to use. However, they are equivalent to the following recurrence relations:

$$
\begin{aligned}
&A_n(s,t) = B_n K(s,t) - n\int_a^{nb-(n-1)a} K(s,u)A_{n-1}(u,t)du,\\
&B_n = \int_a^{nb-(n-1)a} A_{n-1}(t,t)dt.
\end{aligned}
\tag{8}
$$

To give a clear overview of the method of this section, we present a simple example to illustrate the technique. And we can realize the complexity of the computations, too.

Example 1: Consider the following system of Volterra integral equation:

$$
x_1(s) - \int_0^1 tx_2(s)dt = -\frac{s}{2}, \quad x_2(s) - \int_0^1 sx_1(s)dt = s + \frac{3}{2}. \tag{9}
$$

Here,

$$
\lambda = 1,\ a = 0,\ b = 1,\ n = 2.
$$

Rewrite equations (9) as the form:

$$
X(s) - \int_0^2 K(s,t)X(t)dt = Y(s), \tag{10}
$$

where

$$
\begin{aligned}
&X(s) = (x_1(s), x_2(s-1))^T,\\
&Y(s) = (y_1(s),\ y_2(s-1))^T = (-\tfrac{s}{2},\ s + \tfrac{1}{2})^T,\\
&K(s,t) = \begin{pmatrix} k_{11}(s,t) & k_{12}(s,t-1) \\ k_{21}(s-1,t) & k_{22}(s-1,t-1) \end{pmatrix} = \begin{pmatrix} 0 & s \\ t & 0 \end{pmatrix}.
\end{aligned}
$$

Then, from (7), we can get:

$$
\begin{aligned}
&B_0 = 1,\\
&B_1 = \textstyle\int_0^2 K(s,s)ds = \int_0^1 k_{11}(s,s)ds + \int_1^2 k_{12}(s,s-1)ds = 0;\\
&B_2 = \int_0^2\int_0^2 \begin{vmatrix} k_{11}(s,s) & k_{12}(s,t-1)\\ k_{21}(t-1,s) & k_{22}(t-1,t-1)\end{vmatrix} dsdt\\
&\quad = \int_0^1\int_0^1 \begin{vmatrix} k_{11}(s,s) & k_{12}(s,t-1)\\ k_{21}(t-1,s) & k_{22}(t-1,t-1)\end{vmatrix} dsdt\\
&\quad + \int_0^1\int_1^2 \begin{vmatrix} k_{11}(s,s) & k_{12}(s,t-1)\\ k_{21}(t-1,s) & k_{22}(t-1,t-1)\end{vmatrix} dsdt\\
&\quad + \int_1^2\int_0^1 \begin{vmatrix} k_{11}(s,s) & k_{12}(s,t-1)\\ k_{21}(t-1,s) & k_{22}(t-1,t-1)\end{vmatrix} dsdt\\
&\quad + \int_1^2\int_1^2 \begin{vmatrix} k_{11}(s,s) & k_{12}(s,t-1)\\ k_{21}(t-1,s) & k_{22}(t-1,t-1)\end{vmatrix} dsdt\\
&\quad = -\tfrac{2}{3},\\
&B_3 = B_4 = \cdots = 0,
\end{aligned}
$$

So, from (5), the Fredholm determinant

$$d(\lambda) = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!}\lambda^i B_i = B_0 - B_1 + \frac{1}{2!}B_2 = 1 + 0 - \frac{1}{3} = \frac{2}{3}.$$

from (6), we can obtain the first Fredholm minor

$$D_\lambda(s,t) = \frac{1}{3}.$$

So, with formula (3), we have:

$$X(s) = Y(s) + \int_0^2 \frac{D_\lambda(s,t)}{d(\lambda)}Y(s)dt = Y(s) + 2\int_0^2 Y(s)dt,$$

when $0 \le s \le 1$,

$$x_1(s) = -\frac{s}{2} + 2\int_0^1 (-\frac{s}{2})dt + 2\int_1^2 (s - \frac{1}{3})dt = s,$$

when $1 \le s \le 2$,

$$x_2(s) = s - \frac{1}{3} + 2\int_0^1 (-\frac{s}{2})dt + 2\int_1^2 (s - \frac{1}{3})dt = s + 1,$$

All in all, the solution to (9) is as:

$$x_1(s) = s, \quad x_2(s) = s + 1. \tag{11}$$

From this simple example, we know that the iterated collocation method also requires a huge size of calculations. In the following section, we'll establish a new mechanical algorithm for solving system of Fredholm integral equation (1).

3 A New Mechanical Algorithm

The rapid development of computer science and computer algebra system has a profound effect on the concept and the methods of mathematical researches [14,15]. It is a new development orientation in the field of mathematics and computer to conduct science calculation by computer.

The subject, mathematics mechanization, tries to deal with mathematics in a constructive and algorithmic manner so that the reasonings become mechanical, automated, and as much as possible to be intelligence-lacking, with the result of lessening the painstaking heavy brain-labor [14]. Mathematics mechanization is one of important methods of mathematics studies. The basis of mathematics mechanization are algorithm establishing and programming techniques [15,16,17,18,19,20,21,22].

Formulae (3)–(7) can be well adapted to calculate the approximate solution of system of Fredholm integral equation by applying `Maple`. The whole process of the iterated collocation method can be programmed in `Maple`, an exchanged computer algebra system [15]. Now if we want to solve system of Fredholm integral equation (1) based on resolvent method, everything we have to do is just to input the information about the system, then the program will give out the approximate solution of the problem. The main algorithm of `FredEqns` is as follows:

```
FredEqns:=proc(expr::list)
local Expr, parme, i, N, result, num, lambd, TKList, K, xfunction,
    yfunction,bounds,d\_lamb,D\_lamb,m,n,inter,phi,j,kk,Temp;
 result:=0; num:= nops(expr);
 Expr:=map(\_u->Union(\_u),expr);
 for i from 1 to num do parme[i]:=parmeint(Expr[i],2); end do:
 lambd:=parme[1][1];
 TKList:=[seq(map(unapply,parme[i][2],s,t),i=1..num)];
 K:=Matrix(num,2,TKList); bounds:=parme[1][5];
 inter:=bounds[2]-bounds[1];
 xfunction:=map(unapply,[seq(parme[i][3],i=1..num)],s);
 yfunction:=map(unapply,[seq(parme[i][4],i=1..num)],s);
 for m from 1 to num do
    for n from 1 to num do
       K[m,n]:=K[m,n](s-(m-1)*(bounds[2]-bounds[1]),t-(n-1)
              *(bounds[2]-bounds[1]));
       K[m,n]:=unapply(K[m,n],s,t);
    end do:
 end do:
 d\_lamb:=1; D\_lamb:=0; phi:=[seq(0,i=1..num)];
 for i from 1 to num do
```

```
    for j from 1 to num do
        d\_lamb:=1; D\_lamb:=0;
        for kk from 1 to N do
            Temp:=KDetInt(num,kk,K,bounds,i,j);
            d\_lamb:=d\_lamb+lambd\^kk*Temp[1];
            D\_lamb:=D\_lamb+lambd\^kk*Temp[2];
        end do:
        phi[i]:=phi[i]+int(D\_lamb/d\_lamb*yfunction[j](s[2]
                -(j-1)*inter), s[2]=bounds[1]+(j-1)*inter..
                bounds[2]+(j-1)*inter);
    end do:
    phi[i]:=phi[i]+yfunction[i](s[1]-(i-1)*inter);
    phi[i]:=unapply(phi[i],s[1]);
    phi[i]:=simplify(phi[i](s+(i-1)*inter));
end do:
xfunction:=seq(xfunction[i](s-(i-1)*inter),i=1..num);
yfunction:=seq(yfunction[i](s-(i-1)*inter),i=1..num);
lprint('The system can be rewritten as');
print(X(s)-lambda*int('K(x,s)'*X(t),t=bounds[1]..bounds[2]
      +(num-1)*inter)=Y(s));
lprint('where');
print('X(s)'=xfunction,'Y(s)'=yfunction,'K(s,t)'
            =map(\_u->\_u(s,t),K));
lprint('The Fredholm determinant of the kernel K(s,t) is as:');
print('d(lambda)'=d\_lamb);
lprint('and the first order Fredholm subdeterminant is:');
print('D[lambda](s,t)'=subs({s[1]=s,s[2]=t},D\_lamb));
lprint('So, we can get the solution of the equations as:');
print((seq(parme[i][3],i=1..num))=seq(phi[i],i=1..num));
```

Here, the parameter `expr` is the system of Fredholm equation to be solved. For instance, to solve the equations (1) (Example 1) with **FredEqns** in `Maple`, everything we have to do is just to input information about the system as follows:

```
expr[1]:=x[1](s)-int(t*x[2](t),t=0..1)=-s/2;
expr[2]:=x[2](s)-int(t*x[1](t),t=0..1)=s+3/2;
FredEqns([expr[1], expr[2]]);
```

Then we can get the solution as (11).

In the following section, two examples are presented to illustrate the efficiency and accuracy of the mechanized process.

4 Examples

Example 2: Solve the system of Fredholm integral equation:

$$\begin{cases} x_1(s) - \int_0^1 (s+t)x_1(t)dt - \int_0^1 s x_2(t)dt = s^2 - \frac{1}{4} - \frac{7}{3}\,s, \\ x_2(s) - \int_0^1 (s^2+t)x_1(t)dt - \int_0^1 (2\,s+2\,t)x_2(t)dt = -2\,s - \frac{19}{12} - \frac{1}{3}\,s^2. \end{cases} \tag{12}$$

In Maple, with the commands:

```
exprs[3]:=x[1](s)-int((s+t)*x[1](t),t=0..1)-int((s)*x[2](t),
      t=0..1)=s^2-1/4-7/3*s;
exprs[4]:=x[2](s)-int((s2+t)*x[1](t),t=0..1)-int((2*s+2*t)
      *x[2](t),t=0..1)=-2*s-19/12-1/3*s^2;
FredEqns([exprs[3],exprs[4]]);
```

We can get the results immediately:

The system can be rewritten as:

$$X(s) - \lambda \int_0^2 K(s,t)X(s)dt = Y(s)$$

where

$$X(s) = (x_1(s), x_2(s-1)),$$

$$Y(s) = (s^2 - \tfrac{1}{4} - \tfrac{7}{3}\,s, -2\,s + \tfrac{5}{12} - \tfrac{1}{3}\,(s-1)^2),$$

$$K(s,t) = \begin{bmatrix} s+t & s \\ (s-1)^2 + t & 2s - 4 + 2t \end{bmatrix}$$

The Fredholm determinant of the kernel K(s,t) is as:

$$d(\lambda) = -\frac{37}{72}$$

and the first order Fredholm subdeterminant is:

$$D_\lambda(s,t) = \frac{8}{3}\,s - \frac{43}{72} + \frac{10}{9}\,t - \frac{7}{6}\,s^2 - \frac{8}{3}\,st + \frac{7}{6}\,ts^2$$

So, we can get the solution of the equation as:

$$(x_1(s), x_2(s)) = (s^2, 1 + 2s)$$

Example 3: Consider the system:

$$\begin{cases} x_1(s) - \int_0^1 (s+t)\,x_1(t)\,dt - \int_0^1 s x_2(t)\,dt = -\frac{1}{2} + e^s - 3\,es + s, \\ x_2(s) - \int_0^1 (s+t)\,x_1(t)\,dt - \int_0^1 (2\,s + 2\,t)\,x_2(t)\,dt = 2\,e^s - \frac{11}{2} - 5\,es + 2\,s. \end{cases} \tag{13}$$

In **Maple**, one just input the commands:

```
exprs[5]:=x[1](s)-int((s+t)*x[1](t),t=0..1)-int((s)*x[2](t),
      t=0..1)=-1/2+exp(s)-3*exp(1)*s+s;
exprs[6]:=x[2](s)-int((s+t)*x[1](t),t=0..1)-int((2*s+2*t)*x[2](t),
      t=0..1)=2*exp(s)-11/2-5*exp(1)*s+2*s;
FredEqns([exprs[5],exprs[6]]);
```

We can obtain:
`The system can be rewritten as:`

$$X(s) - \lambda \int_0^2 K(s,t)X(s)dt = Y(s)$$

`where`

$$X(s) = (x_1(s), x_2(s-1)),$$

$$Y(s) = (\tfrac{1}{2} + e^s - 3\,es + s, 2\,e^{s-1} - \tfrac{15}{2} - 5\,e\,(s-1) + 2\,s),$$

$$K(s,t) = \begin{bmatrix} s+t & s \\ s-1+t & 2s-4+2t \end{bmatrix}$$

`The Fredholm determinant of the kernel K(s,t) is as:`

$$d(\lambda) = -\frac{11}{18}$$

`and the first order Fredholm subdeterminant is:`

$$D_\lambda(s,t) = -\frac{5}{6}s + \frac{73}{36} - \frac{17}{12}t + \frac{5}{6}st$$

`So, we can get the solution of the equation as:`

$$(x_1(s), x_2(s)) = (1 + e^s, 1 + 2e^s)$$

Example 4: Solve

$$\begin{cases} x_1(s) - \int_0^1 (s+t)x_1(t)dt - \int_0^1 sx_2(t)dt = \sin(s) + \frac{1}{2} - 2\,s + \cos(1)s \\ \qquad - \sin(1) + \cos(1) - 2\,s(\sin(1) + 4), \\ x_2(s) - \int_0^1 (s^2+t)x_1(t)dt - \int_0^1 (2\,s + 2\,t)x_2(t)dt = 2\,\cos(s) + \frac{7}{2} - 2\,s^2 \\ \qquad + \cos(1)s^2 - 5\,\sin(1) - 3\,\cos(1) - 4\,\sin(1)s - 16\,s. \end{cases} \tag{14}$$

With the commands:

```
exprs[7]:=x[1](s)-int((s+t)*x[1](t),t=0..1)-int((s)*x[2](t),t=0
    ..1)=sin(s)+1/2-2*s+cos(1)*s-sin(1)+cos(1)-2*s*(sin(1)+4);
exprs[8]:=x[2](s)-int((s^2+t)*x[1](t),t=0..1)-int((2*s+2*t)
    *x[2](t),t=0..1)=2*cos(s)+7/2-2*s^2+cos(1)*s^2-5*sin(1)
    -3*cos(1)-4*sin(1)*s-16*s;
FredEqns([exprs[7],exprs[8]]);
```

We can have the following results:
`The system can be rewritten as:`

$$X(s) - \lambda \int_0^2 K(s,t)X(s)dt = Y(s)$$

where

$$X(s) = (x_1(s), x_2(s-1)),$$
$$Y(s) = (\sin(s) + \tfrac{1}{2} - 2\,s + \cos(1)s - \sin(1) + \cos(1) - 2\,s(\sin(1)+4),$$
$$2\,\cos(s-1) + \tfrac{39}{2} - 2\,(s-1)^2 + \cos(1)(s-1)^2 - 5\,\sin(1)$$
$$-3\,\cos(1) - 4\,\sin(1)(s-1) - 16\,s),$$
$$K(s,t) = \begin{bmatrix} s+t & s \\ (s-1)^2 + t & 2s - 4 + 2t \end{bmatrix}$$

The Fredholm determinant of the kernel K(s,t) is as:

$$d(\lambda) = -\frac{37}{12}$$

and the first order Fredholm subdeterminant is:

$$D_\lambda(s,t) = \frac{8}{3}s - \frac{43}{72} + \frac{10}{9}t - \frac{7}{6}s^2 - \frac{8}{3}st + \frac{7}{6}ts^2$$

So, we can get the solution of the equation as:

$$(x_1(s), x_2(s)) = (1 + \sin(s), 8 + 2\cos(s))$$

5 Conclusion and Remarks

The resolvent method and a new mechanical algorithm for solving system of Fredholm integral equation have been proposed in this study. The procedure `FredEqns` can give out the approximate solution of system of Fredholm integral equation. The method is very simple and effective for most of the first and second kind of Volterra integral equations.

In the main procedure `FredEqns`, for the sake of reading, we established the form of readable outputting for the results by using the commands `lprint` and `print`. So, the whole processes, solving system of Fredholm integral equation we obtained by `FredEqns`, are just like we do it with papers and our hands. The form of readable outputting may be a good method for computer-aided proving or solving.

References

1. Kythe, P.K., Puri, P.: Computational methods for linear integral equations. Springer, New York (2002)
2. Shen, Y.: Integral equation, 2nd edn. Beijing institute of technology press, Beijing (2002) (in Chinese)
3. Zhang, S.: Integral equation. Chongqing press, Chongqing (1987) (in Chinese)

4. Abdou, M.A.: On the solution of linear and nonlinear integral equation. Appl. Math. Comp. 146, 857–871 (2003)
5. Babolian, E., Biazar, J., Vahidi, A.R.: The decomposition method applied to systems of fredholm integral equations of the second kind. Appl. Math. Comp. 148, 443–452 (2004)
6. Huang, S.C., Shaw, R.P.: The Trefftz method as an integral equation. Adva. in Engi. Soft. 24, 57–63 (1995)
7. Jiang, S.: Second kind integral equations for the classical potential theory on open surface II. J. Comp. Phys. 195, 1–16 (2004)
8. Liang, D., Zhang, B.: Numerical analysis of graded mesh methods for a class of second kind integral equations on real line. J. Math. Anal. Appl. 294, 482–502 (2004)
9. Maleknejad, K., Karami, M.: Using the WPG method for solving integral equations of the second kind. Appl. Math. Comp. 166, 123–130 (2005)
10. Maleknejad, K., Karami, M.: Numerical solution of non-linear fredholm integral equations by using multiwavelets in the Petrov-Galerkin method. Appl. Math. Comp. 168, 102–110 (2005)
11. Maleknejad, K., Mahmoudi, Y.: Numerical solution of linear fredholm integral equations by using hybrid taylor and block-pulse functions. Appl. Math. Comp. 149, 799–806 (2004)
12. Ren, Y., Zhang, B., Qiao, H.: A simple taylor-series expansion method for a class of second kind integral equations. J.Comp. Appl. Math. 110, 15–24 (1999)
13. Yang, S.A.: An investigation into integral equation methods involving nearly sigular kernels for acoustic scattering. J. Sonu. Vibr. 234, 225–239 (2000)
14. Wu, W.: Mathematics mechanization. Science press and Kluwer academic publishers, Beijing (2000)
15. Wang, W.: Computer algebra system and symbolic computation. Gansu Sci-Tech press, Lanzhou (2004) (in Chinese)
16. Wang, W., Lin, C.: A new algorithm for integral of trigonometric functions with mechanization. Appl. Math. Comp. 164, 71–82 (2005)
17. Wang, W.: An algorithm for solving DAEs with mechanization. Appl. Math. Comp. 167, 1350–1372 (2005)
18. Wang, W.: An algorithm for solving nonlinear singular perturbation problems with mechanization. Appl. Math. Comp. 169, 995–1002 (2005)
19. Wang, W., Lian, X.: Computations of multi-resultant with mechanization. Appl. Math. Comp. 170, 237–257 (2005)
20. Wang, W., Lian, X.: A new algorithm for symbolic integral with application. Appl. Math. Comp. 162, 949–968 (2005)
21. Wang, W.: An algorithm for solving the high-order nonlinear volterra-fredholm integro-differential equation with mechanization. Appl. Math. Comp. 172, 1–23 (2006)
22. Wang, W.: Mechanical algorithm for solving the second kind of Volterra integral equation. Appl. Math. Comp. 173, 1149–1162 (2006)

Controllability of Semilinear Impulsive Differential Equations with Nonlocal Conditions

Meili Li[1], Chunhai Kou[1], and Yongrui Duan[2]

[1] Donghua University, Shanghai 201620, PRC
[2] Tongji University, Shanghai, 200092, PRC

Abstract. In this paper we examine the controllability problems of certain impulsive differential equations with nonlocal conditions. Using the Schaefer fixed-point theorem we obtain sufficient conditions for controllability and we give an application.

1 Introduction

In this paper, we consider the following semilinear impulsive differential equation:

$$\begin{aligned} &x'(t) = Ax(t) + Bu(t) + f(t, x(t)), \quad t \in J = [0, b], \quad t \neq t_k, \\ &\Delta x|_{t=t_k} = I_k(x(t_k^-)), \quad k = 1, \cdots, m, \\ &x(0) + g(x) = x_0 \end{aligned} \tag{1}$$

where the state variable $x(\cdot)$ takes values in the Banach space X with the norm $\|\cdot\|$. The control function $u(\cdot)$ is given in $L^2(J, U)$, a Banach space of admissible control functions with U as a Banach space. A is the infinitesimal generator of a strongly continuous semigroup of bounded linear operators $T(t)$ in X, B is a bounded linear operator from U into X. $x_0 \in X$ is the initial state of the system, $0 < t_k < t_{k+1} < t_{m+1} = b < \infty, k = 1, 2, \cdots, m$. $f \in C(J \times X, X), I_k \in C(X, X)(k = 1, \cdots, m)$, $\Delta x|_{t=t_k} = x(t_k^+) - x(t_k^-), x(t_k^+)$ and $x(t_k^-)$ represent the left and right limits of $x(t)$ at $t = t_k$, respectively. $g : PC(J, X) \to X$ are given functions.

The work in nonlocal initial value problems (IVP for short) was initiated by Byszewski[1]. In [1] Byszewski using the method of semigroups and the Banach fixed point theorem proved the existence and uniquess of mild, strong and classical solution of first order IVP. For the importance of nonlocal conditions in different fields, the interesting reader is referred to [1] and the references cited therein. In the past few years, several papers have been devoted to studying the existence of solutions for differential equations with nonlocal conditions[2,3].

Unlike all the previous papers, we will concentrate on the case with impulsive effects. Many evolution processes in nature are characterized by the fact that at certain moments of time an abrupt change of state is experienced. That is the reason for the rapid development of the theory of impulsive differential equations[4].

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 755–762, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The purpose of this paper is to study the controllability of the semilinear impulsive differential equations with nonlocal conditions.

2 Preliminaries

We need the following fixed-point theorem due to Schaefer [5].

Schaefer Theorem. Let S be a convex subset of a normed linear space E and $0 \in S$. Let $F : S \to S$ be a completely continuous operator and let

$$\zeta(F) = \{x \in S : \ x = \lambda Fx \text{ for some } 0 < \lambda < 1\}.$$

Then either $\zeta(F)$ is unbounded or F has a fixed point.

Denote $J_0 = [0, t_1], J_k = (t_k, t_{k+1}], k = 1, 2, \cdots, m$. We define the following classes of functions:

$PC(J, X) = \{x : J \to X : x(t)$ is continuous everywhere except for some t_k at which $x(t_k^-)$ and $x(t_k^+)$ exist and $x(t_k^-) = x(t_k), k = 1, \cdots, m\}$.

For $x \in PC(J, X)$, take $\|x\|_{PC} = \sup_{t \in J} \|x(t)\|$, then $PC(J, X)$ is a Banach space.

Definition 2.1. A solution $x(\cdot) \in PC(J, X)$ is said to be a mild solution of (1) if $x(0) + g(x) = x_0$; $\Delta x|_{t=t_k} = I_k(x(t_k^-)), k = 1, \cdots, m$; the restriction of $x(\cdot)$ to the interval $J_k (k = 0, \cdots, m)$ is continuous and the following integral equation is verified:

$$\begin{aligned} x(t) = {} & T(t)x_0 - T(t)g(x) + \int_0^t T(t-s)[(Bu)(s) + f(s, x(s))]ds \\ & + \sum_{0 < t_k < t} T(t - t_k) I_k(x(t_k^-)), \quad t \in J. \end{aligned} \tag{2}$$

Definition 2.2. The system (1) is said to be nonlocally controllable on the interval J if for every $x_0, x_1 \in X$, there exists a control $u \in L^2(J, U)$ such that the mild solution $x(\cdot)$ of (1) satisfies $x(b) + g(x) = x_1$.

We assume the following hypotheses:

(i) A is the infinitesimal generator of a compact semigroup of bounded linear operators $T(t)$ in X such that $\|T(t)\| \le M_1$, for some $M_1 \ge 1$.

(ii) The linear operator $W : L^2(J, U) \to X$, defined by

$$Wu = \int_0^b T(b - s)Bu(s)ds,$$

has an invertible operator W^{-1}, which takes values in $L^2(J, U)/KerW$ and there exist positive constants M_2, M_3 such that $\|B\| \le M_2$ and $\|W^{-1}\| \le M_3$.

(iii) For almost all $t \in J$, the function $f(t,\cdot) : PC(J,X) \to X$ is continuous and for each $x \in PC(J,X)$, the function $f(\cdot,x) : J \to X$ is strongly measurable.

(iv) For every positive integer k, there exists $\alpha_k \in L^1(J,R_+)$ such that

$$\sup_{\|x(t)\|\le k} \|f(t,x(t))\| \le \alpha_k(t), \text{ a.e. } t \in J,\ x \in PC(J,X).$$

(v) $I_k \in C(X,X)$ and there exist constants d_k such that $\|I_k(x)\| \le d_k$, $k = 1,\cdots,m$.

(vi) $g : PC(J,X) \to X$ is a continuous function and there exists a constant G such that

$$\|g(x)\| \le G, \text{ for } x \in PC(J,X).$$

(vii) There exists an integrable function $m : J \to [0,\infty)$ such that

$$\|f(t,x(t))\| \le m(t)\Omega(\|x(t)\|), \text{ a.e. } t \in J,\ x \in PC(J,X),$$

where $\Omega : [0,\infty) \to (0,\infty)$ is a continuous nondecreasing function with

$$M_1 \int_0^b m(s)ds < \int_c^\infty \frac{ds}{\Omega(s)},$$

where

$$c = M_1(\|x_0\| + G) + M_1 N b + M_1 \sum_{k=1}^m d_k,$$

and

$$N = M_2 M_3[\|x_1\| + G + M_1(\|x_0\| + G) + M_1 \int_0^b m(s)\Omega(\|x(s)\|)ds + M_1 \sum_{k=1}^m d_k].$$

3 Main Result

Theorem 3.1. If the hypotheses (i)-(vii) are satisfied, then system (1) is nonlocally controllable on J.

Proof (of theorem). Using hypothesis (ii) for an arbitrary function $x(\cdot) \in PC(J,X)$, define the control

$$\begin{aligned} u(t) = W^{-1}[x_1 - g(x) - T(b)(x_0 - g(x)) - \textstyle\int_0^b T(b-s)f(s,x(s))ds \\ - \sum_{k=1}^m T(b-t_k)I_k(x(t_k^-))](t). \end{aligned}$$

We shall now show that when using this control the operator defined by

$$\begin{aligned} Fx(t) = T(t)x_0 - T(t)g(x) + \textstyle\int_0^t T(t-s)[(Bu)(s) + f(s,x(s))]ds \\ + \sum_{0<t_k<t} T(t-t_k)I_k(x(t_k^-)),\ t \in J \end{aligned}$$

has a fixed point. This fixed point is then a solution of equation (2).

Clearly, $(Fx)(b) = x_1 - g(x)$, which means that the control u steers the impulsive differential system from the initial state x_0 to x_1 in time b, provided we can obtain a fixed point of the nonlinear operator F.

First we obtain a priori bounds for the following equation:

$$\begin{aligned} x(t) = {} & \lambda T(t)x_0 - \lambda T(t)g(x) + \lambda \textstyle\int_0^t T(t-\eta)BW^{-1}[x_1 - g(x) - T(b) \\ & (x_0 - g(x)) - \textstyle\int_0^b T(b-s)f(s,x(s))ds - \sum_{k=1}^{m} T(b-t_k)I_k(x(t_k^-))](\eta)d\eta \\ & + \lambda \textstyle\int_0^t T(t-s)f(s,x(s))ds + \lambda \sum_{0<t_k<t} T(t-t_k)I_k(x(t_k^-)), \quad t \in J. \end{aligned}$$

We have

$$\begin{aligned} \|x(t)\| & \le M_1(\|x_0\| + G) + M_1 \textstyle\int_0^t M_2M_3[\|x_1\| + G + M_1(\|x_0\| + G) \\ & \quad + M_1 \textstyle\int_0^b m(s)\Omega(\|x(s)\|)ds + M_1 \sum_{k=1}^{m} d_k]d\eta \\ & \quad + M_1 \textstyle\int_0^t m(s)\Omega(\|x(s)\|)ds + M_1 \sum_{k=1}^{m} d_k \\ & \le M_1(\|x_0\| + G) + M_1Nb + M_1 \textstyle\int_0^t m(s)\Omega(\|x(s)\|)ds + M_1 \sum_{k=1}^{m} d_k. \end{aligned}$$

Let $\mu(t) = \sup\{\|x(s)\|, \quad 0 \le s \le t\}(0 \le t \le b)$, then the function $\mu(t)$ is nondecreasing in J, and we have

$$\mu(t) \le M_1(\|x_0\| + G) + M_1Nb + M_1 \int_0^t m(s)\Omega(\mu(s))ds + M_1 \sum_{k=1}^{m} d_k.$$

Denoting by $v(t)$ the right-hand side of the above inequality we have

$$v(0) = M_1(\|x_0\| + G) + M_1Nb + M_1 \sum_{k=1}^{m} d_k \equiv c, \quad \mu(t) \le v(t), \ t \in J,$$

and

$$v'(t) = M_1m(t)\Omega(\mu(t)) \le M_1m(t)\Omega(v(t)), \ t \in J.$$

This implies

$$\int_{v(0)}^{v(t)} \frac{ds}{\Omega(s)} \le M_1 \int_0^b m(s)ds < \int_c^{\infty} \frac{ds}{\Omega(s)}, \ t \in J.$$

This inequality implies that there is a constant K such that $v(t) \leq K (t \in J)$, and hence $\|x(t)\| \leq \mu(t) \leq v(t) \leq K (t \in J)$. So $\|x\|_{PC} \leq K$, where K depends only on b and on the functions $m(\cdot)$ and $\Omega(\cdot)$.

Second, we must prove that the operator $F : PC(J, X) \to PC(J, X)$ is a completely continuous operator.

Let $B_k = \{x(\cdot) \in PC(J, X) : \|x\|_{PC} \leq k\}$ for some $k \geq 1$. We first show that F maps B_k into an equicontinuous family. Let $x \in B_k$ and $t, \bar{t} \in J_k$. Then for $t \leq \bar{t}$, we have

$$
\begin{aligned}
&\|(Fx)(t) - (Fx)(\bar{t})\| \\
&\leq \|T(t)x_0 - T(\bar{t})x_0\| + \|T(t)g(x) - T(\bar{t})g(x)\| \\
&\quad + \| \textstyle\int_0^t [T(t-\eta) - T(\bar{t}-\eta)]BW^{-1} \\
&\quad \cdot [x_1 - g(x) - T(b)(x_0 - g(x)) - \textstyle\int_0^b T(b-s)f(s, x(s))ds \\
&\quad - \textstyle\sum_{k=1}^{m} T(b-t_k)I_k(x(t_k^-))](\eta)d\eta\| \\
&\quad + \| \textstyle\int_t^{\bar{t}} T(\bar{t}-\eta)BW^{-1} \\
&\quad \cdot [x_1 - g(x) - T(b)(x_0 - g(x)) - \textstyle\int_0^b T(b-s)f(s, x(s))ds \\
&\quad - \textstyle\sum_{k=1}^{m} T(b-t_k)I_k(x(t_k^-))](\eta)d\eta\| \\
&\quad + \| \textstyle\int_0^t [T(t-s) - T(\bar{t}-s)]f(s, x(s))ds\| + \| \textstyle\int_t^{\bar{t}} T(\bar{t}-s)f(s, x(s))ds\| \\
&\quad + \| \textstyle\sum_{0<t_k<t} [T(t-t_k) - T(\bar{t}-t_k)]I_k(x(t_k^-))\| \\
&\quad + \| \textstyle\sum_{t<t_k<\bar{t}} T(\bar{t}-t_k)I_k(x(t_k^-))\| \\
&\leq \|T(t) - T(\bar{t})\|(\|x_0\| + G) + \textstyle\int_0^t \|T(t-\eta) - T(\bar{t}-\eta)\|M_2M_3 \\
&\quad \cdot [\|x_1\| + G + M_1(\|x_0\| + G) + M_1 \textstyle\int_0^b \alpha_k(s)ds + M_1 \sum_{k=1}^{m} d_k]d\eta \\
&\quad + \textstyle\int_t^{\bar{t}} \|T(\bar{t}-\eta)\|M_2M_3 \\
&\quad \cdot [\|x_1\| + G + M_1(\|x_0\| + G) + M_1 \textstyle\int_0^b \alpha_k(s)ds + M_1 \sum_{k=1}^{m} d_k]d\eta \\
&\quad + \textstyle\int_0^t \|T(t-s) - T(\bar{t}-s)\|\alpha_k(s)ds + \int_t^{\bar{t}} \|T(\bar{t}-s)\|\alpha_k(s)ds \\
&\quad + \textstyle\sum_{0<t_k<t} \|T(t-t_k) - T(\bar{t}-t_k)\|d_k + \sum_{t<t_k<\bar{t}} \|T(\bar{t}-t_k)\|d_k.
\end{aligned}
$$

The right-hand side is independent of $x \in B_k$ and tends to zero as $\bar{t} - t \to 0$, since the compactness of $T(t)$ for $t > 0$ implies the continuity in the uniform operator topology.

Thus, the functions in FB_k are equicontinuous on $J_k (k = 0, 1, \cdots, m)$.

It is easy to see that the family FB_k is uniformly bounded.

Next, we show $\overline{FB_k}$ is compact. Since we have shown FB_k is equicontinuous collection, it suffices by the Arzela-Ascoli theorem to show that F maps B_k into a precompact set in X.

Let $0 < t \leq b$ be fixed and ϵ a real number satisfying $0 < \epsilon < t$. For $x(\cdot) \in B_k$, we define

$$\begin{aligned}
&(F_\epsilon x)(t)\\
&= T(t)x_0 - T(t)g(x) + \int_0^{t-\epsilon} T(t-\eta)BW^{-1}\\
&\quad \cdot[x_1 - g(x) - T(b)(x_0 - g(x))\\
&\quad - \int_0^b T(b-s)f(s,x(s))ds - \sum_{k=1}^{m} T(b-t_k)I_k(x(t_k^-))](\eta)d\eta\\
&\quad + \int_0^{t-\epsilon} T(t-s)f(s,x(s))ds + \sum_{0<t_k<t} T(t-t_k)I_k(x(t_k^-))\\
&= T(t)x_0 - T(t)g(x) + T(\epsilon)\int_0^{t-\epsilon} T(t-\eta-\epsilon)BW^{-1}\\
&\quad \cdot[x_1 - g(x) - T(b)(x_0 - g(x))\\
&\quad - \int_0^b T(b-s)f(s,x(s))ds - \sum_{k=1}^{m} T(b-t_k)I_k(x(t_k^-))](\eta)d\eta\\
&\quad + T(\epsilon)\int_0^{t-\epsilon} T(t-s-\epsilon)f(s,x(s))ds + \sum_{0<t_k<t} T(t-t_k)I_k(x(t_k^-)).
\end{aligned}$$

Since $T(t)(t > 0)$ is a compact operator, the set $X_\epsilon(t) = \{(F_\epsilon x)(t) : x \in B_k\}$ is precompact in X, for every $\epsilon, 0 < \epsilon < t$. Moreover, for every $x(\cdot) \in B_k$, we have

$$\begin{aligned}
&\|(Fx)(t) - (F_\epsilon x)(t)\|\\
&\leq \int_{t-\epsilon}^{t} \|T(t-\eta)BW^{-1}[x_1 - g(x) - T(b)(x_0 - g(x))\\
&\quad - \int_0^b T(b-s)f(s,x(s))ds - \sum_{k=1}^{m} T(b-t_k)I_k(x(t_k^-))](\eta)\|d\eta\\
&\quad + \int_{t-\epsilon}^{t} \|T(t-s)f(s,x(s))\|ds\\
&\leq \int_{t-\epsilon}^{t} \|T(t-\eta)\|M_2M_3[\|x_1\| + G + M_1(\|x_0\| + G) + M_1\int_0^b \alpha_k(s)ds\\
&\quad + M_1\sum_{k=1}^{m} d_k]d\eta + \int_{t-\epsilon}^{t} \|T(t-s)\|\alpha_k(s)ds.
\end{aligned}$$

Therefore, there are precompact sets arbitrarily close to the set $\{(Fx)(t) : x \in B_k\}$. Hence,the set $\{(Fx)(t) : x \in B_k\}$ is precompact in X.

It is easy to see that $F : PC(J,X) \to PC(J,X)$ is continuous. This completes the proof that F is completely continuous.

Finally, the set $\zeta(F) = \{x(\cdot) \in PC(J,X) : x = \lambda Fx, \lambda \in (0,1)\}$ is bounded, as we proved in the first step. Consequently, by Schaefer Theorem, the operator F has a fixed point in $PC(J,X)$. This means that any fixed point of F satisfying $(Fx)(t) = x(t)$ is a mild solution of problem (1) on J. Thus, the system (1) is nonlocally controllable on J. □

4 An Application

More precisely we consider the following nonlocal IVP

$$\begin{aligned}
&x'(t) = A(t)x(t) + (Bu)(t) + \int_0^t l(t,s)f(s,x(s))ds,\ t \in J = [0,b],\ t \neq t_k,\\
&\Delta x|_{t=t_k} = I_k(x(t_k^-)),\ k = 1,\cdots,m,\\
&x(0) + g(x) = x_0,
\end{aligned} \tag{3}$$

where $A(t)$ is a linear closed densely defined operator in a Banach space X, and $f : J \times X \to X,\ l : J \times J \to R,\ g : PC(J,X) \to X$ are given functions. Also, the control function $u(\cdot)$ is given in $L^2(J,U)$ and B is a bounded linear operator from U to X, where U is a Banach space and $L^2(J,U)$ is a Banach space of admissible control function with norm $\|u\|_{L^2} = (\int_0^b \|u(t)\|^2 dt)^{1/2}$.

It is known that the IVP (3) can be written as a nonlinear integral equation

$$\begin{aligned} x(t) &= U(t,0)x_0 - U(t,0)g(x) \\ &\quad + \int_0^t U(t,s)[(Bu)(s) + \int_0^s l(s,\tau)f(\tau,x(\tau))d\tau]ds \\ &\quad + \sum_{0<t_k<t} U(t,t_k)I_k(x(t_k^-)),\ t \in J, \end{aligned}$$

where $\{U(t,s) : 0 \le s \le t \le b\}$ is a strongly continuous family of evolution operator on X. We shall make the following assumptions on the evolution system $U(t,s)$:

(U_1) $U(t,s) \in L(X)$, the space of bounded linear transformations on X, whenever $0 \le s \le t \le b$ and for each $x \in X$ the mapping $(t,s) \to U(t,s)x$ is continuous;

(U_2) $U(t,s)U(s,r) = U(t,r),\ 0 \le r \le s \le t \le b$;

(U_3) $U(t,t) = I$, the identity operator on X;

(U_4) $U(t,s)$ is a compact operator whenever $t - s > 0$.

Sufficient conditions for $(U_1) - (U_4)$ to hold may be found in Friedman [6].

For the study of this problem we list the following hypotheses.

(H_1) The linear operator $W : L^2(J,U) \to X$, defined by

$$Wu = \int_0^b U(b,s)Bu(s)ds,$$

has an invertible operator W^{-1}, which takes values in $L^2(J,U)/KerW$ and there exist positive constants M_2, M_3 such that $\|B\| \le M_2$ and $\|W^{-1}\| \le M_3$.

(H_2) $l : J \times J \to R$ is a continuous function, and there exists a constant L such that

$$|l(t,s)| \le L,\ t \ge s \ge 0.$$

(H_3) $\hat{M}L \int_0^b \int_0^t m(s)dsdt < \int_{c_1}^{\infty} \frac{ds}{\Omega(s)}$,
where

$$M = \sup\{\|U(t,s)\| : 0 \le s \le t \le b\},\ c_1 = \hat{M}(\|x_0\| + G) + \hat{M}N_1 b,\ \hat{M} = \max\{1, M\},$$

$$N_1 = M_2M_3[\|x_1\| + G + M(\|x_0\| + G) + ML\int_0^b \int_0^s m(\tau)\Omega(\|x(\tau)\|)d\tau ds + M\sum_{k=1}^{m} d_k].$$

Theorem 4.1. Let $\{U(t,s) : 0 \le s \le t \le b\}$ satisfying $(U_1) - (U_4)$. Assume also that the hypotheses (iii)-(vi),$(H_1) - (H_3)$ hold. Then the IVP (3) is nonlocally controllable on J.

Proof (of theorem). The proof is similar to that of Theorem 3.1, so we omit it. □

5 Conclusion

The controllability of infinite-dimensional systems in Banach spaces has been studied extensively by virtue of the fixed point arguments. The essential part of this method is to transform the controllability problem into a fixed point problem for an appropriate operator in a function space. By introducing the new concepts such as mild solution and nonlocal controllability of impulsive delay differential equations, we establish some sufficient conditions for the controllability of the system (1). An example of impulsive delay partial differential equations is presented to illustrate the theory.

References

1. Byszewski, L.: Theorems about the existence and uniqueness of solutions of a semilinear evolution nonlocal Cauchy problem. J. Math. Anal. Appl. 162, 494–505 (1991)
2. Balachandrany, K., Ilamaran, S.: Existence and uniqueness of mild and strong solutions of a semilinear evolution with nonlocal condition. Indian J. Pure. Appl. Math. 25, 411–418 (1994)
3. Balachandrany, K., Chandrasekaran, M.: Existence of solutions of a delay-differential equation with nonlocal conditions. Indian J. Pure. Appl. Math. 27, 443–449 (1996)
4. Bainov, D.D., Simeonov, P.S.: Systems with Impulse Effect. Ellis Horwood, Chichester (1989)
5. Schaefer, H.: Über die Methode der a priori Schranken. Mathematische Annalem 129, 415–416 (1955)
6. Friedman, A.: Partial Differential Equations. Holt, Rinehart & Winston, New york (1969)

Dynamic Adaptation of Workflow Based Service Compositions

Heiko Pfeffer, David Linner, and Stephan Steglich

Technische Universität Berlin
Franklinstr. 28/29, 10587 Berlin, Germany
{heiko.pfeffer,david.linner,stephan.steglich}@tu-berlin.de
http://www.tu-berlin.de

Abstract. Service composition mechanisms successfully enable 'programming in the large' within business oriented computing systems. Here, the composability of single software components provides dynamicity and flexibility in the design of large-scale applications. However, this dynamicity is restricted to the late binding of services to service interface descriptions; the workflow, i.e. the execution order of the single services, remains static.

Within this paper, we present the modification of abstract service composition plans, extending service compositions' dynamicity from late binding of service implementations to a dynamic reconfiguration of the service composition structure itself. As a proof of concept, we demonstrate the adaptability of the service composition plan by applying genetic operators on the workflow graph of the service composition. Finally, an evaluation mechanism is presented to estimate the degree of similarity between the resulting composition plan and present ones.

1 Introduction

Service composition, i.e. the loosely coupled combination of single services to a large-scale application, constitutes a key pillar of the success of Service Oriented Architectures (SOA). By enabling reusability of software components, flexibility in application configuration and collaboration among different parties to achieve higher-level goals, the principle of service composition is the main enabler to meet today's SOA requirements. Here, the dynamicity provided by service composition relies on the composition of interface descriptions instead of service implementations, rendering possible a late binding of the actual services. Semantic extensions of those service descriptions provide means to address services based on their functionality and non-functional properties, abstracting from the concrete service interface during service discovery.

However, today's approaches to increase dynamicity in service composition procedures focus on the dynamic aggregation of services only; the service composition plan itself,i.e. the specification of the execution order of and interaction among single single services remains static.

Within this paper, we built on a service composition model constituting of two control graphs specifying the service compositions workflow and data flow

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 763–774, 2008.
© Springer-Verlag Berlin Heidelberg 2008

between single services, respectively [1]. We show that this model builds an appropriate basis for adapting service compositions in their structural representation. Therefore, we exemplarily outline the transformation of service composition plans through the application of genetic operators and discuss their subsequent evaluation.

2 State of the Art and Related Work

SOA and WebServices together with their various models for describing large-scale business processes are the success story of modern service composition approaches [2]. The most prominent language is the Business Process Modeling Language (BPEL), originating from the conflation of Microsoft's block-oriented XLANG [3] and IBM's graph-based WebService Flow Language (WSFL) [4]. BPEL has become a OASIS WebService standard in 200t, referred to as WS-BPEL version 2.0, formerly BPEL4WS [5,6]. However, there are various other composition languages, feraturing advantages and disadvantages within different use cases; van der Aalst et al. thus assent that there isn't yet an holistically accepted standard for WebService Composition [7]. For instance, ebXML containing the Business Process Specification Schema (BPSS) [8] represents an alternative composition language that is rarely used within Europe and the US, but is common in Asia.

Multiple semantic descriptions such as DAML-s [9] and OWL-S [10] have been proposed within the scope of Semantic Web activities, enabling the dynamic discovery of WebServices and automatic generation of compositions. Especially for mobile devices, more lightweight descriptions have been proposed in order to reduce the computation complexity entailed by semantic reasoning processes [11]. However, the automatic creation of appropriate service descriptions for automatically generated service compositions is still challenging. Within this paper, we thus aim at modifying existing service compositions in order to produce equivalent or similar service compositions with regard to their functionality, which may however differ in their non-functional properties. Those modifications of service compositions are exemplarily performed through the application of genetic operators. Canfora et al. [12] have used a similar approach for the QoS-aware creation of service compositions. However, they did not focus on the problem of related semantic service descriptions.

3 Modeling Service Compositions

Within this section, we introduce service composition model based on two control graphs connecting semantically annotated service blueprints. Therefore, we introduce a service model featuring semantically extended I/O descriptions in section 3.1 and discuss their composition in section 3.2. The formal definition of the according control graphs is finally outlined in section 3.3, summarizing the main definitions presented in [1].

3.1 Semantically Enhanced Service Model

Services themselves are considered as atomically executable parts of application logic, whereby the execution of the service is independent from outer computations and data structures. Instead of relying on IOPE descriptions characterizing a service functionality by its inputs, outputs, preconditions and effects as discussed by Jaeger et al. [13], the presented service model relies on extended I/O descriptions. The I/O descriptions are initially restricted to primitive data types to ease the transformation of service compositions. Inputs and outputs of a service are distinguishable by a unique identifier referred to as a *port*.

As an extension of the classic I/O description patterns, services may generate a special type of semantically described outputs, referred to as *gains*, defining what a user/requester 'gains' by executing the service. Those gains are assumed to possess an importance beyond simple I/O passage within the service composition and thereby considerably constitute the services functionality; they are globally managed within a *gain space*, where they are accessible for all services within a service composition.

For instance, a Restaurant Finder service may get a String as input, specifying the current location of the user by an address within a city. Assume the output of the service is a Boolean, indicating whether a restaurant was found in, e.g., a radius of 2 kilometers. In case a restaurant was found, a map is pushed to the requesting user indicating the location of the restaurant. Within the presented service model, such a map is specified as a semantically described gain. Notably, this gain is a key feature of the service, specifying its functionality more profoundly than the simple I/O pattern.

Figure 1 shows an exemplary service designed with regard to the introduced service model.

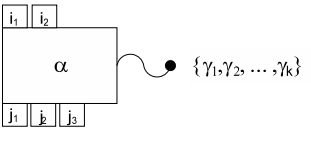


Fig. 1. Service Model

It requires two inputs i_1, i_2 for executing the atomic action α, entailing the generation of three outputs j_1, j_2, j_3 and s gains $\gamma_1, \gamma_2, ..., \gamma_k$ with $s \leq k$.

Within the subsequent section, we describe how this service model can serve as grounding for the description of complex service compositions.

3.2 Service Composition Model

In order to represent service compositions properly as well as to control their execution, two main concepts are introduced in the following. First, we describe the control of the service compositions by a bipartite graph concept; afterwards, we

discuss the global handling of gains within a service composition. For representing the execution order of services, interleaving as well as their possible parallel execution, a workflow graph is introduced controlling the execution of the single components within a service composition. Each transition of the graph corresponds to a service, which is modeled as introduced in section 3.1, i.e. contains I/O parameters and a list of gains it can generate during execution.

When passing a transition from one vertice of the workflow graph to another one, an action representing a service functionality is performed, which consumes the service's input and generates a finite set of gains and a finite set of outputs. Edges between vertices are optionally annotated with guards, which restrict the transition of the edge by requiring the presence of specific gains or previously generated outputs. Since outputs of a service are not necessarily required as inputs for a service reached by the next edge but may become relevant after multiple other services have been executed, a second graph is defined, specifying the dataflow within a service composition. This dataflow graph represents the flow of outputs from one service component to the input of another with its edges. The passage of a transition can thus be restricted by a guard operating on a set of gains and on the availability of all inputs that have to be created as outputs of other services before.

As motivated, gains generated during the execution of services are semantically described and supposed to represent the functionality encapsulated by the services. For instance, a service receiving a string as input and representing a map on a screen showing the location of the city specified within the input string may have a Boolean as output, indicating whether the service execution was successful or not. The coevally generated gain is an abstraction of the fact that the map is pushed to the user in case the service execution was successful. Notably, (as outlined in section 3.1) only a subset of the annotated gains has to be generated during the service execution. Given other inputs, the same service may perform a direct call setup between the requester and the restaurant nearby. This call setup would also be annotated to the service as a gain. Which of the two gains is really generated then depends on the inputs and cannot be derived from the service description itself.

Within the approach presented in this work, gains are stored globally within a gain space. The gain space thus enables a global view of gains generated during the execution of a service composition. Access restrictions for the gain space emerging from security related issues have to be formulated by special guards; however, this topic is out of scope for this paper and part of currently ongoing research. Beside the semantic representation of a gain itself, the service that has generated it is kept. By handling gains globally during the execution of a service composition, transition guards within the workflow graph can operate on all gains already generated. Thus, a guard may check whether a specific gain has been created or whether it has been created by the appropriate service component.

In the following section, the bipartite graph representation of a service composition is formally introduced, specifying its guard based transition behavior and its gain consideration mechanisms.

3.3 Bipartite Graph Concept

The representation of the workflow graph bases on timed automata and has already been formally introduced in our previous work [1]. In the following, we will shortly outline the most important parts of the graph representation of service compositions.

In general, a timed automaton [14,15] is finite automaton extended by a set of real-time valued clocks; for the remainder of this paper, we follow the definition provided by Clarke et al. [16]. In short, a timed automaton is a digraph with nodes that are connected with transitions. A transition itself is labeled with a guard and an action. In case the guard is satisfied, the passage of a transition leads to the execution of the respective action. For timed automata, guards can also operate on a special type of variables, so-called *clocks*, which are real-valued variables that progress simultaneously. Thus, in case one clock is increased by a value δ, all other clocks are augmented by the same δ. Moreover, a *reset* operation can be applied to a finite set of clocks, resetting their value to zero when passing an according transition.

Since actions are supposed to model the behaviour of a service, the execution of an action α is mapped to the consumption of inputs, the creation of outputs and the generations of gains. Inputs and outputs are regarded as typed variables bounded to specific ports, enabling the definition of I/O passage by means of a dataflow graph. The domains of variables thus constitute their primitive data type.

We model Workflow Graphs Θ as timed automata, extended by the ability of actions to operate on typed variables and gains. Typed variables are defined as 2-tuples $(x, \Gamma(x))$, where x is a variable and $\Gamma(x)$ its respective domain. For now, we only consider variables x with $x \in R \subset \mathbb{R}$. Let $\mathcal{V}$ be a finite set of typed variables. Regarding the gain space $\mathcal{G}$, we define $\mathcal{G}^*$ as a final set of variables $\gamma^* \in \{0, 1\}$ indicating whether a gain γ has been created (γ^* set to 1) or not (γ^* set to 0). We assume that $\mathcal{V}$ contains at least all variables of the gain space, thus, $\mathcal{G}^* \subseteq \mathcal{V}$.

Guards are logical formulas evaluating to a Boolean, blocking the passage of a transition within the workflow graph in case it is evaluated to *false*. Out model operates on *guard constraints* $\mathcal{C}(X \circ R \backslash \mathcal{V})$ that contain *clock constraints* $\mathcal{C}(X)$ and *real-value constraints* $\mathcal{C}(R \backslash \mathcal{V})$ over $\mathcal{V}$. A clock constraint is a conjunctive formula of atomic constraints of the form $x * n$ or $x - y * n$, where x and y are considered as real-valued clocks, $* \in \{\leq, <, =, >, \geq\}$ and $n \in \mathbb{N}$. A guard constraint is a propositional logic formula $x * n$, where $x \in \mathcal{V}$ is a typed variable, $n \in \mathbb{R}$ and $* \in \{<, \leq, \geq, >, ==\}$. A guard constraint is either a clock constraint, a real-value-constraint, or a conjunctive form of both.

We now define the workflow graph as a timed automaton whose actions and guards can also operate on typed variables [1].

Definition 1 (Workflow Graph). *A workflow graph Θ is a timed automaton, where*

1. *Σ is a finite alphabet. The alphabet represents the actions which are identified by the single services within the service composition.*

2. *S is a finite set of locations defining the service compositions's current state,*
3. *$s_0 \in S$ are the initial locations defining the initial state of the service composition,*
4. *X is a set of clocks,*
5. *$I : S \rightarrow \mathcal{C}(X)$ assigns invariants to locations, restricting the system in the amount of time it is allowed to remain in the current state, and*
6. *$T \subseteq S \times \mathcal{C}(X \circ R \backslash \mathcal{V}) \times \Sigma \times 2^X \times S$ is the set of transitions, denoting the execution of a service (represented by an action α). The passage of a transition (and thus the execution of a service) thereby depends on whether the according guard is met.*

Abbreviatory, we will write $s \overset{g,a,\lambda}{\rightarrow} s'$ for $\langle s, g, \alpha, \lambda, s' \rangle$, i.e. the transition leading from location s to s'. The transition is restricted by the constraint g (often called guard); $\lambda \subseteq X$ denotes the set of clocks that is reset during the transition passage.

A formal specification of the transition relation of the proposed workflow graph can be found in our previous work [1]; we conclude the key principle of the theory in the following. The state transition relation has to enable two basic operations: the resetting of clocks to zero and the execution of an action. Therefore, two transition relations are introduced. The *delayed transition* enables a timed automaton the remain in a state and elapse a specific amount of time as long as the state invariant is not violated. The *action transition* describes the passage of a transition from a state s to a state s', preconditioning that the transition guard is fulfilled and the state invariant of s' is not violated. During the passage of the transition, a finite amount of time can be elapsed. However, this passage of time is optionally and restricted by the state invariant of s'.

Figure 2 shows an exemplary part of a workflow graph and its respective transition annotations.

In order to decide, whether an output of a service component is required as an input for another one, a dataflow graph is introduced in the following [1]. Dataflow graphs keep the same locations as the workflow graph introduced in Definition 1 and use labeled transitions to express inter-component data passing.

Definition 2 (Dataflow Graph Ω). *A dataflow graph Ω is a labeled directed digraph $\Omega = \{N, P, E\}$, where,*

1. *N is a final set of labeled nodes, which is equivalent to the set of locations S held in the according workflow graph Θ,*
2. *P is a set of port mappings represented by 2-tuples $p = (p_1, p_2)$ indicating the passage of the output from port p_1 to the input port p_2, and*
3. *$E \subseteq N \times P \times N$ is a final set of labeled directed transitions.*

Each location represents the data generated by the execution of action α_i; we therefore label a node with $d(\alpha_i)$ indicating the output data from action α_i. If a transition is passed within a workflow graph Θ entailing the execution of action α_i, all outgoing transitions from location $d(\alpha_i)$ within the according dataflow graph Ω are passed. A transition passage $d(\alpha_i) \overset{(p_m,p_n)}{\rightarrow} d(\alpha_j)$ within Ω effectuates

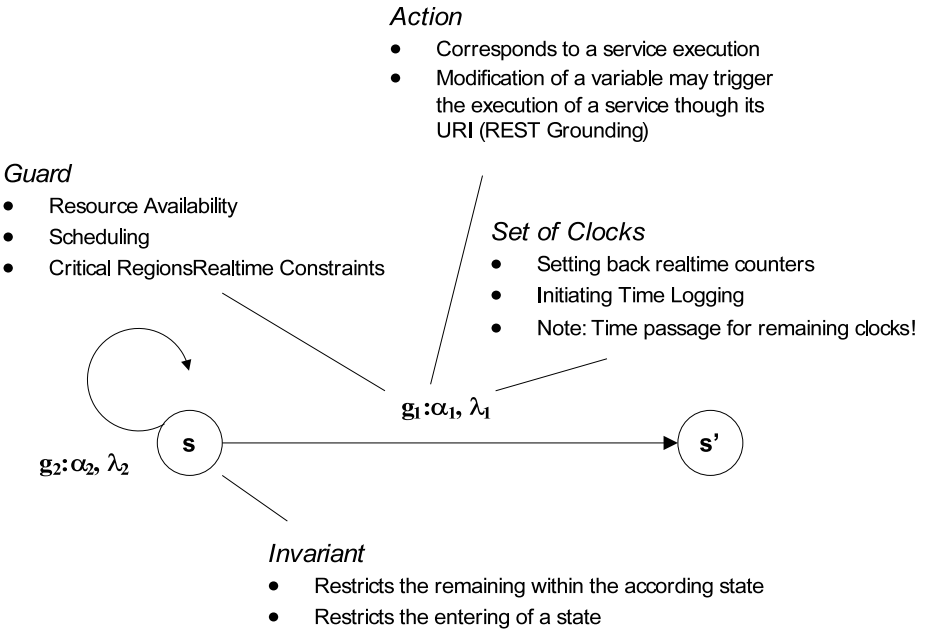


Fig. 2. Abstract Fragment of a Workflow Graph

that the output at port p_m from action α_i is redirected to input port p_n of action α_j in case action α_i is executed within Θ.

A service composition is defined by its workflow and dataflow graph, i.e. is defined as a 2-tuple $\langle \Theta, \Omega \rangle$.

4 Applying Genetic Operators to Service Compositions

As a proof of concept for the dynamicity of the introduced service composition model, we present techniques for the generation of new service compositions through the application of generic operators. Therefore, existing service compositions are transformed and subsequently evaluated with regard to their functional equivalence to other present service compositions. Within section 4.1, a definition for stable service compositions is introduced, specifying the class of proper service composition transformations. Section 4.2 then discusses the genetic operators that can be applied to service compositions in order to modify their structure. Section 5 later deals with the evaluation of stable service compositions with regard to their functionality.

4.1 Stable Service Composition Transformations

A *Service Composition Transformation* is defined as the swapping of a non-empty set of transitions within the workflow graph of a service composition.

Thus, the set of transitions $T = \{t_0, ..., t_i, ...t_{i+m}, ..., t_k\}$ is replaced by $T' = \{t_0, ..., t'_i, ...t'_{i+n}, ..., t_k\}$. The swapping is not restricted to sets of transitions of the same cardinality. Thereby, inputs of some services may not be available any more, since they have been produced by actions modeled as transitions that have been removed by the transformation procedure. The workflow and dataflow graph both have to be restructured in order to integrate the new transitions and the related I/O flow. In case every input of the new dataflow graph is generated -enabling its execution with regard to holistic input availability- , the service composition is regarded as *stable* and can be evaluated by means of its functionality as discussed in section 5. Otherwise the composition is considered to be morbid and is therefore discarded. Within the next section, two genetic operators are introduced, defining the transformation of the workflow and dataflow graph by exchanging transitions as well as their final test of stability.

4.2 Genetic Operators

Service Composition aims at creating added-value by reusing existing functional elements. In addition to service discovery, service selection, and late-binding, decentralized and dynamic computing environments require to consider varying availability of services. Service compositions are not bindable if for at least one of the constituting services there is no instance available.

For this reason, we investigate mechanisms for the availability-aware modification of service compositions on a structural level while preserving the semantics of the resulting services. In [17] we presented techniques for the generation of service composition variations through the appliance of generic algorithms. Existing service compositions are transformed and subsequently evaluated with regard to their functional equivalence to other service compositions. Therefore, the straightforward applicability of genetic operators was one of the major requirements for the definition of the graph presented in this work.

For the beginning, we realized a cross-over operator and a mutation operator. Both operators are basically built on inclusion and removal of graph elements. Usually, a transformation on the workflow graph also requires the dataflow graph to be recovered, while a transformation on the dataflow graph has no impact on the consistency of the workflow graph. Since both operators, mutation and cross-over are not restricted in the way they modify a workflow graph, there is a marginal probability that the corresponding dataflow graph will not be recoverable afterwards. In this case, the generated offspring is rejected as candidate before evaluation.

The mutation operator is randomly applied to workflow graph or dataflow graph. A transformation on the workflow is implemented to either remove transitions or include new ones by random. The inclusion of transitions depends on the types of services available in the service environment. The removal of transitions requires the subsequent recovery of the dataflow graph. If the mutation operator is only applied to the dataflow graph, it randomly changes the passages of data between input ports and output ports, removes constants, or includes new ones, all by preserving the correct typing.

The cross-over operator is applied to the workflow graph only. The operator also removes and includes transitions, except for the difference that the removed transitions are exchanged with the workflow graph of another service composition, while the transitions obtained in return are included. Afterwards, the dataflow graphs of both compositions are recovered with respect to the newly introduced transitions.

First practical tests for service composition transformation by applying mutation and cross-over operator showed the appropriateness of the graph representation presented in this work for flexible modification in dynamic environments.

5 Evaluation of Functional Equivalence of Service Compositions

By the transformation of service compositions as introduced in section 4, executable service compositions are created whose functionality is unknown. Research effort has been spent on identifying the functionality of a service and to create an appropriate service description accordingly [18]. However, the dynamic creation of semantic service description is still an open problem when desisting from approaches basing on constrictive assumptions as within [19]. Currently, our work aims at transferring semantic descriptions instead of creating new ones. Therefore, the new description is geared to the description of the service is was derived from by transformation. The new service is thus evaluated with regard to its similarity in its functional behavior to other available service compositions.

Since the generation of gains during the execution of a service may depend on the actual service inputs, services cannot be evaluated offline with regard to their functionality (which is considerably dependent from the gains it produces), but have to be estimated during runtime. Therefore, newly created service compositions are compared with existing ones in a *Silent Execution Mode*. Here, two compositions are executed within a Sandbox [8] like environment, where their execution does not directly affect the current state of the computing environment. During silent execution, initial inputs are randomly generated and a selected service composition as well as the transformed one are executed with those input parameters. The outputs and generated gains of both service compositions can then be compared by a distance function δ allowing to make a statement on the proximity of effect and outputs the service compositions have been provided. A possible distance function δ can be given as follows.

Definition 3 (Distance Function δ). *Let m be the number of outputs, n the number of gains created by at least one of two service compositions. During a silent execution run, the distance $\delta \in (0,1)$ between the compositions is given as*

$$\delta = \frac{\sum_{i=0}^{m-1} outAccur_i + \sum_{j=0}^{n-1} gainAccur_j}{m+n},$$

where $outAccur_i$ *and* $gainAccur_j \in (0,1)$ *denote the level of similarity of output* i *or gain* j *of both compositions, respectively.* $gainAccur_j$ *is 1 in case gain* j *has been created by both service composition executions, else 0. The value of* $outAccur_i$ *depends on the primitive type of the output variable. For* `String` *and* `Char`*, a test of equivalence is performed, i.e.* $outAccur_i = 1$ *if the* `Char` *or* `String` *created by the compositions match exactly, otherwise 0. For numeric variables such as* `Int`*,* `Float` *and* `Double`*,* $outAccur_i$ *is given by* $\frac{out_{i1}}{out_{i2}}$ *where* out_{i1} *is the smaller of the two outputs. In case an output was only created by one composition,* $outAccur_i$ *is 0. If multiple runs are performed for calculating the distance of two service compositions,* δ *is given as the arithmetic average of the single distance values.*

A higher similarity between both compositions is thus expressed through a higher value of δ. We identify the deviance in terms of output and gain accuracy between two service compositions with the deviance in service composition functionality. That is, two service compositions generating exactly the same outputs and gains are considered as functionally equivalent. The distance δ moreover remains a variation parameter for service composition evaluation. For instance, a composition may be considered as equivalent if the distance to another service composition is bigger than 0.9, because the deviance may be caused by simulation inaccuracy, different service contexts or randomness.

The exactness of such an evaluation technique apparently also depends on the number of simulation runs used to calculate the distance between the two service compositions' outputs. The number of simulation runs therefore constitutes the second parameter (beside the acceptance border for δ) enabling the assignment of a certain level of trust to a simulated distance between two service compositions. For instance, two service compositions may be regarded as functionally equivalent, if they have a distance greater than 0.8 which was derived from at least 50 simulation runs. In case a new service composition c' derived by the appliance of a genetic operator to a composition c is evaluated as functionally equivalent to c, the semantic description of c is overtaken for c'.

6 Conclusion and Future Prospect

Within this paper, we presented a workflow model for the representation of service compositions featuring rapid modifiability through gain-oriented service descriptions. The transformation of service composition plans has been exemplarily demonstrated by the application of algorithms derived from genetic programming. Finally, a method for the evaluation of resulting service compositions has been outlined, enabling the comparison of service composition plans with regard to the outputs and gains they generate.

The automatic creation of workflow oriented service composition plans based on a requested set of gains is part of ongoing research.

References

1. Pfeffer, H., Linner, D., Steglich, S.: Modeling and Controling Dynamic Service Compositions. In: Proceedings of The Third International Multi-Conference on Computing in the Global Information Technology (ICCGI 2008), Athens, Greece (2008)
2. Bucchiarone, A., Gnesi, S.: A Survey on Services Composition Languages and Models. In: Bertolino, A., Polini, A. (eds.) Proceedings of International Workshop on Web Services Modeling and Testing (WS-MaTe2006), Palermo, Sicily, Italy, pp. 51–63 (2006)
3. Thatte, S.: XLANG - Web Services for Business Process Design (2001)
4. Leymann, F.: Web Service Flow Language (WSFL 1.0). In: IBM (May 2001)
5. Andrews, T., Curbera, F., Dholakia, H., Goland, Y., Klein, J., Leymann, F., Liu, K., Roller, D., Smith, D., Thatte, S., Trickovic, I., Weerawarana, S.: Business Process Execution Language for Web Services, Version 1.1 (May 2003), (Accessed, November 2007), `http://www.ibm.com/developerworks/library/specification/ws-bpel/`
6. Curbera, F., Goland, Y., Klein, J., Leymann, F., Roller, D., Thatte, S., Weerawarana, S.: Business Process Execution Language for Web Services (Version 1.0) (July 2002)
7. van der Aalst, W.: Don't go with the flow: Web services composition standards exposed (2003)
8. UNCEFACT and OASIS: ebXML Business Process Specification Schema Version 1.0.1 (2001)
9. Ankolekar, A., Burstein, M., Hobbs, J.R., Lassila, O., Martin, D., McDermott, D., McIlraith, S.A., Narayanan, S., Paolucci, M., Payne, T., Sycara, K.: DAML-S: Web Service Description for the Semantic Web. In: Horrocks, I., Hendler, J. (eds.) ISWC 2002. LNCS, vol. 2342. Springer, Heidelberg (2002)
10. The OWL Services Coalition: OWL-S: Semantic Markup for Web Services (November 2004) (Accessed February 26, 2007), `http://www.daml.org/services/owls/1.1/`
11. Pfeffer, H., Linner, D., Jacob, C., Steglich, S.: Towards Light-weight Semantic Descriptions for Decentralized Service-oriented Systems. In: Proceedings of the 1st IEEE International Conference on Semantic Computing (ICSC 2007). Volume CD-ROM, Irvine, California, USA (2007)
12. Canfora, G., Di Penta, M., Esposito, R., Villani, M.L.: An approach for QoS-aware service composition based on genetic algorithms. In: GECCO 2005: Proceedings of the 2005 Conference on Genetic and Evolutionary Computation, pp. 1069–1075. ACM, New York (2005)
13. Jaeger, M., Engel, L., Geihs, K.: A Methodology for Developing OWL-S Descriptions. In: First International Conference on Interoperability of Enterprise Software and Applications Workshop on Web Services and Interoperability (INTEROP-ESA 2005), Springer, Heidelberg (2005)
14. Alur, R., Courcoubetis, C., Dill, D.: Model-checking for real-time systems. In: Proceedings of of the 5th Annual Symposium on Logic in Computer Science, pp. 414–425. IEEE Computer Society Press, Los Alamitos (1990)
15. Dill, D.: Timing assumptions and verification of finite-state concurrent systems. In: Sifakis, J. (ed.) Proceedings of the International Workshop on Automatic Verification Methods for Finite State Systems. LNCS, vol. 407, pp. 197–212. Springer, Heidelberg (1989)

16. Clarke, E.M.J., Grumberg, O., Peled, D.A.: Model Checking. The MIT Press, Cambridge (1999)
17. Linner, D., Pfeffer, H., Steglich, S.: A genetic algorithm for the adaptation of service compositions. In: Proceedings of the 2nd International Conference on Bio-Inspired Models of Network, Information, and Computing Systems (2007)
18. Babik, M., Hluchy, L., Kitowski, J., Kryza, B.: Generating Semantic Descriptions of Web and Grid Services. In: Kacsuk, P., Fahringer, T., Nemeth, Z. (eds.) Distributed and Parallel Systems - From Cluster to Grid Computing (Proceedings of the 6th Austrian-Hungarian Workshop on Distributed and Parallel Systems (DAPSYS)), Innsbruck, Austria. International Series in Engineering and Computer Science (ISECS), pp. 83–93 (2006)
19. Caprotti, O., Davenport, J.H., Dewar, M., Padget, J.: Mathematics on the (Semantic) NET. In: Bussler, C.J., Davies, J., Fensel, D., Studer, R. (eds.) ESWS 2004. LNCS, vol. 3053, pp. 213–224. Springer, Heidelberg (2004)

The Research and Application of Nonlinear Predictive Functional Control Based on Characteristic Models

Peijian Zhang[1], Jianguo Wu[1, 2], and Minrui Fei[2]

[1] College of Information Science and Engineering, Nantong University, Nantong 226019, China
zhang.pj@ntu.edu.cn
[2] School of Mechatronics and Automation, Shanghai University, Shanghai 200072, China
mrfei@staff.shu.edu.cn

Abstract. By combining dynamical feature of objects with requirements of control performance, predictive model uses a second order slow time varying linear model to characterize original controlled station. This paper introduces a way to establish a category of predictive models of nonlinear objects and lead this kind of model into predictive functional control (PFC) so then obtain nonlinear PFC arithmetic based on characteristic model. We successfully avoid the problem of online identification of slow time varying parameters by setting parameter interval, thereby this kind of control arithmetic is typical of easy operation, high speed and wide applicability in comparison with traditional PFC. Finally this paper offers applied living examples by considering temperature control problem of continuous stirring tank reactor(CSTR).

Keywords: Predictive functional control, characteristic model, non – linear temperature control.

1 Introduction

Among industrial controls, we usually can only achieve effective control when we obtain mathematical model of controlled station. These models can be parametric ones and can also be non parametric ones. Theoretically (or by mechanism modeling) some controlled stations can only be presented in precise term by high order model, but during practical produce process it is too complicated to work out such a model for controlled station, thus scientists always try their best to simplify the model under effective control. Setting up characteristic model for controlled station proposed by academician Wu Hongxing is just on account of this kind of thought. This kind of characteristic model uses a second order slow time varying linear model characterizing original controlled station by combining dynamical characteristic of objects with requirements of control performance. It's of great significant to utilize characteristic model to replace practical complex produce process control. Some basic principles for describing characteristic model are as following:

①With same input object characteristic model and actual objects are equivalent on output (characteristic model can keep outputs in allowable errors in dynamic process), the results are the same under stable condition.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 775–783, 2008.
© Springer-Verlag Berlin Heidelberg 2008

②The form and order of characteristic model mainly depend on requirements of control performance besides consideration of image characteristics.

③Characteristic model function should be simpler, easier and more convenient achieved by engineering; Characteristic model compresses related information from high order model into several characteristic parameters without losing information, which is not the same as reduced order model. In general cases, characteristic model is presented by slow time varying differential equation.

Eq. (1) shows linear time invariant high order object:

$$G(s)=\frac{k_v}{s^2}+\sum_{i=1}^{m}\frac{k_i}{s+\lambda_i}+\sum_{i=1}^{h}\left(\frac{k_{p+i}}{s+\lambda_{p+i}}+\frac{\bar{k}_{p+i}}{s+\lambda_{p+i}}\right)+\sum_{i=1}^{p}\frac{k_{wi}}{(s+w_i)^2}. \tag{1}$$

Inside, $m \le n$ and $n \ge 2$, with a certain sampling period Δt, Eq. (1) can be presented in form of a second order time varying differential equation as following, when we want to realize position – keeping or position – tracking.

$$y(k+1)=f_1(k)y(k)+f_2(k)y(k-1)+g_0(k)u(k)+g_1(k)u(k-1). \tag{2}$$

Eq. (2) is a characteristic model of linear time invariant high order object (1).

Suppose controlled station has time delay, its differential equation can be rewritten as following:

$$y(k+1)=f_1(k)y(k)+f_2(k)y(k-1)+g_0(k)u(k-k_0)+g_1(k)u(k-k_0-1). \tag{3}$$

For produce progress with self – regulating when time – varying is not so serious, Eq. (3) can be transferred into linear time invariant model. Obviously, it is similar with the principles of model order reducing.

The objects from Eq. (1) belong to linear time invariant high order ones. For non linear objects, academician Wu Hongxing also posed corresponding characteristic model subjects. This paper mainly studies on those setting characteristic model of non linear objects and introducing obtained models into predictive functional control arithmetic on the purpose of realizing nonlinear PFC.

2 Characteristic Model of Nonlinear System [4-5]

Setting a series of non-linear system as Eq. (4)

$$\dot{x}(t)=f(x,\dot{x},\cdots,x^{(n)},u,\dot{u},\cdots,u^{(m)})=f(x,u). \tag{4}$$

If do $x=x_1,\dot{x}=x_2,\cdots,x^{(n)}=x_{n+1},u=u_1,\dot{u}=u_2,\cdots,u^{(m)}=u_{m+1}$ then Eq.(4) can be written as:

$$\dot{x}_1(t)=f(x_1,\cdots,x_{n+1},u_1,\cdots,u_{m+1}). \tag{5}$$

Suppose above nonlinear system has following characteristics:

① single input and single output as well as controlled

② the order of control volume $u(t)$ is 1

③ when both independent variables are 0, function $f(\bullet)=0$

④ $f(\bullet)$ is continuous derivative to all independent variables and all partial derivatives are bounded

⑤ $\left|f(x(t+\Delta t),u(t+\Delta t))-f(x(t),u(t))\right|<M\Delta t$, M is a normal value, Δt is a sampling period

⑥ Independent variables x_i, u_i are bounded

For assumption ⑥, it can be met easily because control energy can not be unbounded in engineering, control value $u(t)$ is always bounded thus independent variables like x_i, u_i are the same.

This kind of nonlinear system can be presented by following linear system [6]:

$$\begin{aligned}\dot{x}_1(t)=&\alpha_1(t)x_1(t)+\alpha_2(t)x_2(t)+\cdots+\alpha_{n+1}(t)x_{n+1}(t)\\&+\beta_1(t)u_1(t)+\beta_2(t)u_2(t)+\cdots+\beta_{m+1}(t)u_{m+1}(t)\end{aligned}. \tag{6}$$

On the basis of assumption①, we have

$$\begin{aligned}\dot{x}_1(t)=&f(x_1,\cdots,x_{n+1},u_1,\cdots,u_{m+1})-f(0,0,\cdots,0)\\=&f(x_1,x_2,\cdots,x_{n+1},u_1,u_2,\cdots,u_{m+1})-f(x_1,x_2,\cdots,x_{n+1},u_1,u_2,\cdots,u_m,0)\\&+\cdots+f(x_1,x_2,\cdots,x_{n+1},u_1,0,\cdots,0)-f(x_1,x_2,\cdots,x_{n+1},0,\cdots,0)\\&+\cdots+f(x_1,0,\cdots,0)-f(0,\cdots,0)\end{aligned}. \tag{7}$$

According to assumption ②and mean value theorem:

$$\begin{aligned}&f(x_1,\cdots,x_i,0,\cdots,0,0,\cdots,0)-f(x_1,\cdots,x_{i-1},0,\cdots,0,0,\cdots,0)\\&=\left.\frac{\partial f}{\partial x_i}\right|_{(z_i,r_i,0,\cdots,0)}x_i=\alpha_i(t)x_i \qquad (i=1,2,\cdots,n+1,)\end{aligned}.$$

r_i is a certain number between 0 and x_i

$$\begin{aligned}&f(x_1,\cdots,x_{n+1},u_1,\cdots,u_j,0,\cdots,0)-f(x_1,\cdots,x_{n+1},u_1,\cdots,u_{j-1},0,\cdots,0)\\&=\left.\frac{\partial f}{\partial u_j}\right|_{(z_j',r_j',0,\cdots,0)}u_j=\beta_j(t)u_j \qquad (j=1,2,\cdots,m+1)\end{aligned}.$$

r_j' is a certain number between 0 and u_j

Among, $z_i=(x_1,\cdots,x_i), z_j'=(x_1,\cdots,x_{n+1},u_1,\cdots,u_j,0,\cdots,0)$

Thus, Eq. (6) can be written as following:

$$\dot{x}_1(t)=\alpha_1(x_1)x_1+\beta_1(x_1,\cdots,x_{n+1})u_1+F_x(\bullet)+F_2(\bullet). \tag{8}$$

Among, $\alpha_1(x_1)=\dfrac{\partial f(x_1,0,\cdots,0)}{\partial x_1}$ is a function of x_1,

$$\beta_1(\bullet)=\frac{\partial f(x_1,\cdots,x_{n+1},u_1,0,\cdots,0)}{\partial u_1}$$

As u is a first order parameter. $\beta(\bullet)$ is only a function of $(x_1,\cdots,x_{n+1})$. $F_x(\bullet)$ is a function of x_i $(i=1,2,\cdots,n+1)$, $F_2(\bullet)$ is a function of x_i and u_j $(i=1,2,\cdots,n+1,\ j=1,2,\cdots,m+1)$.

Define $F_x(\bullet)+F_2(\bullet)=F(\bullet)$

Differentiate Eq. (8)

$$\begin{aligned}\ddot{x}_1 &= \frac{d[\alpha_1(x_1)x_1]}{dt}+\frac{d[\beta_1(\bullet)u_1]}{dt}+\frac{d[F(\bullet)]}{dt}\\ &= \frac{\partial[\alpha_1(x_1)x_1}{\partial x_1}\frac{dx_1}{dt}+\frac{d\beta_1(\bullet)}{dt}u_1+\beta_1(\bullet)\frac{du_1}{dt}+\frac{d[F(\bullet)]}{dt}\end{aligned}. \tag{9}$$

Adding Eq. (8) and Eq. (9) together, we can get Eq. (10) by using method of approximate discretion.

$$\begin{aligned}&\frac{x_1(k+1)-2x_1(k)+x_1(k-1)}{(\Delta t)^2}+\frac{x_1(k)-x_1(k-1)}{\Delta t}\\ &=\frac{\partial[\alpha_1(x_1)x_1]}{\partial x_1}\frac{[x_1(k)-x_1(k-1)]}{\Delta t}\\ &\quad+\frac{[\beta_1(k)-\beta_1(k-1)]}{\Delta t}u_1(k)+\beta_1(k)\frac{[u_1(k)-u_1(k-1)]}{\Delta t}+\frac{F(k)-F(k-1)}{\Delta t}\\ &\quad+\alpha_1(k)x_1(k)+\beta_1(k)u_1(k)+F(k)\end{aligned}. \tag{10}$$

$$x_1(k+1)=f_1'(k)x_1(k)+f_2'(k)x_1(k-1)+g_0'(k)u_1(k)+g_1'(k)u_1(k-1)+W(k). \tag{11}$$

Among,

$$f_1'(k)=2-\Delta t+\Delta t[\frac{\partial[\alpha_1(x_1)x_1]}{\partial x_1}]+(\Delta t)^2\alpha_1(k)$$

$$f_2'(k)=-(1-\Delta t+\Delta t[\frac{\partial[\alpha_1(x_1)x_1]}{\partial x_1}])$$

$$g_0'(k)=\Delta t\beta_1(k)+\Delta t(\beta_1(k)-\beta_1(k-1))+(\Delta t)^2\beta_1(k)$$

$$g_1'(k)=-\Delta t\beta_1(k)$$

$$W(k)=\Delta t[F(k)-F(k-1)]+(\Delta t)^2F(k)$$

For standardization, we change Eq. (11) into following:

$$x(k+1)=f_1(k)x(k)+f_2(k)x(k-1)+g_0(k)u(k)+g_1(k)u(k-1). \tag{12}$$

Among,

$$f_1(k)=2-\Delta t+\Delta t[\frac{\partial[\alpha_1(x_1)x_1]}{\partial x_1}]+(\Delta t)^2\alpha_1(k)+\frac{W(k)x_1(k)}{x_1^2(k)+x_1^2(k-1)}$$

$$f_2(k) = -\{1 - \Delta t + \Delta t[\frac{\partial[\alpha_1(x_1)x_1]}{\partial x_1}] - \frac{W(k)x_1(k-1)}{x_1^2(k) + x_1^2(k-1)}\}$$

$$g_0(k) = \Delta t\beta_1(k) + \Delta t(\beta_1(k) - \beta_1(k-1)) + (\Delta t)^2\beta_1(k)$$

$$g_1(k) = -\Delta t\beta_1(k)$$

Eq.(12) and original system are dynamic equivalent while steady state same. As in dynamic process, although $W \neq 0$, $F(k)$ is bounded thus $\Delta t^2 F(k)$ is too small to arouse errors.

In Eq. (12) $f_i(k), g_i(k)$ are both slow time varying as

$$\Delta_1 = \left|\frac{f_i(k) - f_i(k-1)}{f_i(k)}\right|, \Delta_2 = \left|\frac{g_i(k) - g_i(k-1)}{g_i(k)}\right|.$$

While $\Delta t \to 0$, $\Delta_1 \to 0, \Delta_2 \to 0$

Choose Δt properly, while $\Delta_1 < 0.01\%, \Delta_2 < 0.01\%$, system can be considered as slow time varying one.

From above analysis, we know that no matter it is high order linear time invariant system or above two kinds of non linear objects, we can obtain unified second order slow time varying linear model, which makes model become more simple.

As characteristic model is a second order one and its pole can spread the whole S plane that can reflect all kinds of modality, which can more precisely characterize controlled station .Adopting second order characteristic model as predictive model to forecast model control can decrease mismatch between predictive model and controlled station so that we can gain better control results. Meanwhile, introducing characteristic model which reflect nonlinear objects into predictive functional control can bring out non linear predictive functional control in view of characteristic modeling.

3 Nonlinear Predictive Function Control in View of Characteristic Model

Following is a flow chart of predictive functional control (PFC) in brief [7-9]. Applying linear combination of several known base functions $(n = 1, \cdots, N)$ to achieve PFC which is different from traditional predictive control.

$$u(k+i) = \sum_{n=1}^{N} \mu_n f_n(i) \quad i = 0,1 \cdots P-1. \tag{13}$$

Here, P is optimize time – domain length, is coefficient of linear combination. The choice of base functions depend on nature of set value, usually we take base functions like step – function, slope function, index, etc.

The essence of introducing characteristic model into PFC is utilizing characteristic model to structure predictive model of PFC. Predictive model is abstracted from controlled station. The degree of approximation between predictive model and controlled

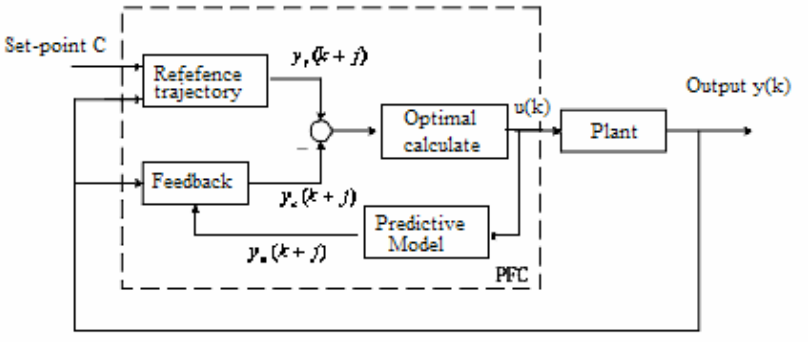


Fig. 1. Flow chart of PFC

station influences performance of control system directly. For PFC, predictive model usually adopts first order model or second order one linear time invariant .In order to use PFC to realize controlling of non linear objects, we use its characteristic model as predictive model of PFC.

Considering unitarily and time lag, we adopt following predictive model.

$$\begin{aligned} y_m(k+1) = {} & f_1(k)y_m(k) + f_2(k)y_m(k-1) + g_0(k)u(k-k_0) \\ & + g_1(k)u(k-k_0-1) \end{aligned} . \tag{14}$$

Assume standard function of minimization as,

$$\min J = \sum_{i=0}^{P-1} (y_r(k+i) - y_m(k+i) - e(k+i))^2 . \tag{15}$$

Here, $y_r(k+i)$ is reference trail for a steady system, usually take first order index form as reference trail.

$$y_r(k+i) = c(k+i) - \beta^i (c(k) - y(k)) . \tag{16}$$

Among, $c(k)$ is track set value, $\beta = e^{(-T_S/T_r)}$, T_s is sampling period . T_r is 95% response time of reference track.

$e(k+i)$ is a predictive of error between model output and practical output. Generally future error is

$$e(k+i) = y(k) - y_m(k) . \tag{17}$$

We can get μ_n by optimizing parameters, to obtain control value $u(k)$ finally.

4 Application Cases

Continuous stirring tank reactor (CSTR) is a common device in chemical industry. Fig.2 is its schematic diagram. Its nonlinear model is

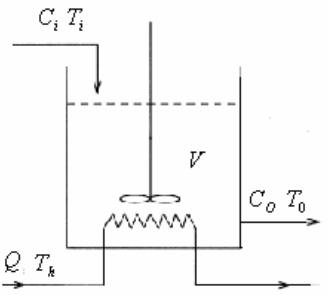


Fig. 2. Schematic diagram of CSTR

$$\dot{C}_o = \frac{q}{V}(C_i - C_o) - a_0 C_o e^{-\frac{E}{RT_o}}. \tag{18}$$

$$\dot{T}_o = \frac{q}{V}(T_i - T_o) + a_1 C_o e^{-\frac{E}{RT_O}} + aQ[1 - e^{-\frac{a_2}{Q}}](T_h - T_o). \tag{19}$$

We can use the method of modeling by testing way to obtain a group of mathematical models. Their transfer functions are:

$$\begin{bmatrix} T_o(s) \\ C_o(s) \end{bmatrix} = \begin{bmatrix} \dfrac{K_{11}e^{-\tau_{11}s}}{(T_{11}S+1)(T_{11}^{'}S+1)} & \dfrac{K_{12}e^{-\tau_{12}s}}{(T_{12}S+1)(T_{12}^{'}S+1)} \\ \dfrac{K_{21}e^{-\tau_{21}s}}{(T_{21}S+1)(T_{21}^{'}S+1)} & \dfrac{K_{22}e^{-\tau_{22}s}}{(T_{22}S+1)(T_{22}^{'}S+1)} \end{bmatrix} \begin{bmatrix} Q(s) \\ C_i(s) \end{bmatrix}. \tag{20}$$

Eq.(20) can easily be transferred into Eq.(14) second order feature model.

From Eq.(20), we see there is coupling between flow (or liquid level, volume) and temperature. We do not decouple for reflecting the advantages of feature model.

In stead of adopting slow time varying parameters for feature functions, we set range of parameter variation on the basis of a group of mathematic model functions from modeling in reality, see Eq.(21).

$$\begin{aligned} f_1 &\in [f_1^-, f_1^+] \\ f_2 &\in [f_2^-, f_2^+] \\ g_0 &\in [g_0^-, g_0^+] \\ g_1 &\in [g_1^-, g_1^+] \end{aligned}. \tag{21}$$

Its advantage is avoiding on line identification of parameters to make control algorithm more simply and faster.

Fig.3 is a control curve from experimental system. Among, line 1 is a manipulated variable of temperature control, line 2 is a set value of temperature, line 3 is an actual testing value of temperature. Line 4 is a variety of liquid level. We can see actual

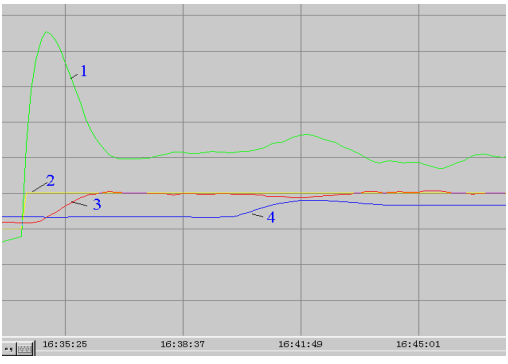


Fig. 3. Experimental Curve

temperature has no overshoot and rises fast during step – function (temperature from 30℃ to 45℃). Increase disturbances (change liquid level) on purpose, temperature are still small. Thus we can conclude that the robustness of PFC algorithm in term of feature model is stronger.

5 Conclusion

For produce process control, nonlinear phenomenon is quite common. For PFC, it is unavailable to realize nonlinear control by designing nonlinear predictive model. This paper established feature models of nonlinear objects and introduced this model into PFC and take advantage of coefficient interval to solve on line parameter disguise problem for varying coefficient.

As the forms of linear time invariant high order systems and normal feature models of common nonlinear systems are always the same, this method can be applied widely. By contrast, the shortcoming of traditional PFC is its predictive model is too simple to characterize controlled station precisely, which will result in worse control performance because of mismatch of models. Applying feature model to establish predictive model of PFC can improve control performance of PFC largely.

Acknowledgments. This project was supported by the Program for Application Study of Nantong (K2007017) and Nantong University Natural Science Fund(06Z038).

References

1. Wu, H.X.: Intelligent Characteristic Model and Intelligent Control. Acta automatica sinica 28 Supp.1, 30–37 (2002)
2. Wu, H.X., Xie, Y.C., Li, Z.B., He, Y.Z.: Intelligent Control based on Description of Plant Characteristic Model. Acta automatica sinica 25(1), 9–17 (1999)
3. Wu, H.X., Wang, Y.C., Xing, Y.: Intelligent Control based on Intelligent Characteristic Model and its Application. SCIENCE IN CHINA (Series E) 32(6), 805–816 (2002)

4. Wu, H.X., Wang, Y., Xie, Y.C.: Nonlinear Golden Section Adaptive Control. Journal of Astronautics 23(6), 1–8 (2002)
5. Wu, H.X., Wang, Y., Xie, Y.C.: Characteristic Modeling and Control Method of Nonlinear System. Control Enginerring 6, 1–7 (2002)
6. Guo, J., Chen, Q.W., Zhu, R.J., He, W.L.: Adaptive Predictive Control of a Class of Nonlinear System. Control theory and Applications 19(1), 68–72 (2002)
7. Richalet, J., Doss, S.A., Arber, C., et al.: Predictive Functional Control: Applications to Fast and Accurate Robots. In: Isermann, R. (ed.) Automatic Control 10th Triennial World Congress of IFAC, pp. 251–258. Pergamon Press, Oxford (1988)
8. Richalet, J., Rault, A., Testud, J.L., Papon, J.: Model Predictive Heuristic Control: Application to Industrial Processes. Automatics 14(5), 413–428 (1978)
9. Wu, J.G., Zhang, P.J.: Study of Pure Time Delay System based on Predictive Functional Control. Process automation instrumentation 25(11), 8–10 (2004)

A Hybrid Particle Swarm Optimization for Manipulator Inverse Kinematics Control

Xiulan Wen, Danghong Sheng, and Jiacai Huang

Automation Department, Nanjing Institute of Technology, Nanjing, China, 211167
zdhxwxl@njit.edu.cn

Abstract. A very important problem usually encountered in the study of robot manipulators is the inverse kinematics problem. The inverse kinematics control of a robotic manipulator requires solving non-linear equations having transcendental functions and involving time-consuming calculations. In this paper, a hybrid particle swarm optimization based on the behaviour of insect swarms and natural selection mechanism is firstly presented to optimize neural network (HPSONN) for manipulator inverse kinematics. Compared with the results of the fast back propagation learning algorithm (FBP), conventional genetic algorithm (GA) based elitist reservation (EGA), improved GA (IGA) and immune evolutionary computation (IEC), the simulation results verify the hybrid particle swarm optimization is more effective for manipulator inverse kinematics control than above most methods.

Keywords: Hybrid Particle Swarm Optimization, Neural Network, Manipulator, Inverse Kinematics Control.

1 Introduction

Manipulator inverse kinematics control can be represented by nonlinear continuous function. And it is a nonlinear and configuration dependent problem that may have multiple solutions [1]. Many methods have been proposed for the inverse transformation problem. There are two basic approaches: closed-form and iterative solutions. Most iterative techniques utilize the manipulator Jacobin J . However, iterative inverse transform solutions using simply J^{-1} will not converge around singularities, since the joint rates tend to infinity as the singularity is approached [2].

The feed forward neural network has been widely applied to the pattern recognition, the servo control, the robot control and so on. A 3-layer feed-forward neural network (FFNN) can approximate any nonlinear continuous function to any arbitrary accuracy [3]. This is especially useful for complex optimization problems where the number of parameters is large and the analytical solutions are difficult to obtain such as solving manipulator inverse kinematics [4, 5]. The traditional learning algorithm of FFNN is back propagation learning algorithm (BP). But it has the problems of local minima and parameter sensitivity [6]. Therefore, Evolutionary algorithms (EAs) are used to train FFNN [4, 5]. EAs, which include GA, Evolutionary strategies (ES), Evolutionary programming (EP) and Genetic programming (GP) are a class of adaptive search techniques

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 784–791, 2008.
© Springer-Verlag Berlin Heidelberg 2008

based on the principles of population evolution. Each individual in the population represents a candidate solution to a concrete problem and has an associated fitness value, which measures the goodness of the solution. EAs have been theoretically and empirically proven to be robust for searching solutions in complex spaces and have been widely used in optimization, training neural networks, estimating parameters in system identification, etc. However, many research results indicate that premature convergence is still the preeminent problem encountered in evolutionary search procedures. Another frequently encountered problem is the relatively low-search efficiency. For solving complex optimization problems, traditional evolutionary computation methods may take a long time to find the solution [7]. The natural immune system is a complex adaptive pattern-recognition system that defends the body from foreign pathogens (bacteria or virus). The clonal selection and affinity maturation principles are used to explain how the immune system reacts to pathogens and how it improves its capability of recognizing and eliminating pathogens. Inspired by the clonal selection, affinity maturation principles and the mutation ideas, an immune evolutionary computation (IEC) is proposed to evolve neural network [8]. The particle swarm optimization (PSO) is similar to evolutionary computational (EC) techniques, in which a population of potential solutions to the problem under consideration is used to probe the search space. The major difference is that the EC techniques use genetic operators whereas swarm intelligence techniques use the physical movements of the individuals inspired by social behavior of a flock of birds and insect swarms. PSO is originally proposed by Kennedy and Eberhart[9]. Due to its simple mechanism and high performance for global optimization, PSO has been applied to many optimization problems successfully [10-11]. However, PSO is potentially premature convergence and sometimes terminates as other EAs. The combination of PSO and a selection mechanism of EC can generate a high-quality within short calculation time [12, 13]. In this paper, a hybrid particle swarm optimization (HPSO) based on the behaviour of insect swarms and natural selection mechanism is presented to evolve neural networks for manipulator inverse kinematics control which is a complex optimization problem. The simulation results verify that the HPSO provides much better generalization performance than FBP, EGA, IGA and slightly better than IEC.

2 A Hybrid Particle Swarm Optimization

In PSO, a swarm is a collection of particles. A particle has both a position and a velocity vector. And each particle flies around in the multidimensional search space. Each dimension of this space represents a parameter of the problem to be optimized. During the search procedure that seeks the best location in the solution domain, the particles change their position with time as they fly around in a multidimensional solution space. During their flight, particles adjust their velocity according to their own experience and those of the other particles, making use of the best position researched, both by themselves and by the rest of the swarm. As a result, the particle is stochastically attracted both toward the individual's best position ever found (*pbests*), as well as toward the group's best position (*gbest*) while evaluating the goodness of its current location in the process. Since PSO finds the optimal solution deeply depending on *pbests* and *gbest* and the searching area is limited by *pbests* and

*gbes*t. HPSO based on the behavior of insect swarms and natural selection mechanism is firstly used to train feed-forward neural networks for manipulator inverse kinematics control. The best particle with the minimum objective function value is considered as the solution to trained neural network.

The flow that HPSO optimizes neural network is as follows:

Step 1: Randomly generate initial position and initial velocity of all particles in whole search space. In applying HPSO to optimize neural network, the weight and the biases of the neural network link are defined as the positions of particles.

Step 2: Calculate the objective function of all particles, which is defined as the mean square error of the output of feed-forward neural network. The smaller the objective function value is, the better the particle is.

Step 3: Select better particles. Particles positions with larger objective function values are replaced by those with smaller objective function values using selection. The replaced rate is called selection rate.

Step 4: Update velocity. The velocity of a particle is determined according to the relative position from *pbest* and *gbest*, and there are two common methods to modify the velocity, that is inertia weights approach (IWA) and constriction factor approach (CFA) [14]. CFA is employed to modify the velocity. Because CFA of PSO ensures the convergence of the search procedures based on the mathematical theory. CFA can generate higher-quality solutions than PSO with IWA [15].

$$v_i^{t+1} = K(v_i^t + C_1 r_1(pbest_i^t - p_i^t) + C_2 r_2(gbest^t - p_i^t)) \tag{1}$$

$$K = \frac{2}{|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}|}, \text{ where } \varphi = c_1 + c_2, \varphi > 4$$

where v_i^t and p_i^t are the velocity and position of *i*th particle at iteration *t*, respectively. r_1 and r_2 are uniform random numbers between 0 and 1. C_1 and C_2 are acceleration factors that determine the relative pull for each particle toward *pbest* and *gbest*. The convergence characteristic of the system can be controlled by φ.

Step5*:* Update position. Each particle modifies its position according to the following equation

$$p_i^{t+1} = p_i^t + v_i^t \Delta t \tag{2}$$

Where $\triangle t$ is a time step, which is usually set be one.

Step6: Update *pbests*. If the current objective function value of a particle is less than the old *pbest* value, the *pbest* is replaced by the current position.

Step 7: Update *gbest*. If the current objective function value of a particle is less than the old *gbest* value, the *gbest* is replaced by the current position.

Step 8: Go to Step 2 until the maximum iteration is satisfied.

3 HPSONN for Manipulator Inverse Kinematic Control

3.1 The Manipulator Inverse Kinematics

Formally, consider a manipulator with *K* degrees of freedom. For a given manipulator, at any instant of time, a joint configuration establishes a unique position and orientation of the end effector in the Cartesin space. At any instant *t*, denote its joint variables by $q = [q_1, q_2, ..., q_K]^T \in R^K$. Also, define the position variables describing the manipulator tasks by a vector of *L* variables $x = [x_1, x_2, ..., x_L]^T \in R^L$, R^K and R^L are the K-dimensional and the L-dimensional Euclidean spaces, respectively. x and q are related by

$$x = f(q) \tag{3}$$

where Eq.(3) is the forward kinematics function of manipulator. The inverse kinematics problem is to find the inverse mapping of Eq. (4)**:**

$$q = f^{-1}(x) \tag{4}$$

The most direct way to solve Eq. (4) is to derive a closed-form solution from Eq. (3), Due to the nonlinearity of $f(\cdot)$, obtaining a closed-form solution is impossible for most robots. Making use of the relation between joint velocity $\dot{q}(t)$ and Cartesian velocity $\dot{x}(t)$ is another approach to the inverse kinematics problem. The velocity vectors $\dot{q}(t)$ and $\dot{x}(t)$ have the following linear relation:

$$\dot{x}(t) = J(q)\dot{q}(t) \tag{5}$$

where $J(q)$ is the $L \times K$ Jacobian matrix and can be singular in general. In order to determine $\boldsymbol{q}(t)$ for given $\boldsymbol{x}(t)$, the velocity vector of the joints $\dot{q}(t)$ needs to be computed. $\dot{q}(t)$ can be computed as:

$$\dot{q}(t) = J^+(q(t))\dot{x}(t) + [I - J^+(q(t))J(q(t))]k(t) \tag{6}$$

where $J^+(q(t))$ is the pseudoinverse of $J(q(t))$, I is an identity matrix, and $k(t)$ is a K-vector of arbitrary time-varying variables. Obviously, the pseudoinverse Jacobian is crucial for the computation of $\dot{q}(t)$. The methods using Eq. (6) have the following disadvantages: a nonsingular $J^+(q(t))$ for all time cannot be entirely guaranteed; and a cyclic behavior cannot be ensured [16].

3.2 Neural Network Model

The position of a particle is set to be the weight and the biases of the neural network link $[\omega_{ji}^{(1)}, \theta_j^{(1)}, \omega_{kj}^{(2)}, \theta_k^{(2)}]$ for all i, j, k, which is an array of floating point

numbers. The 3-layer FFNN is described by

$$y_k = \tilde{\varphi}((\sum_{j=1}^{n_{node}} \omega_{kj}^{(2)} \varphi(\sum_{i=1}^{n_{in}} \omega_{ji}^{(1)} x_i + \theta_j^{(1)}) + \theta_k^{(2)})) \tag{7}$$

where x_i and y_k are the i-th input and the k-th output of the neural network, $i = 1, 2, ..., n_{in}$, $j = 1, 2, ..., n_{node}$, $k = 1, 2, ..., n_{out}$; n_{in}, n_{node} and n_{out} denote the number of input, hidden and output nodes, respectively. $\omega_{ji}^{(1)}$ denotes the weight link between the i-th input node and the j-th hidden node. And $\omega_{kj}^{(2)}$ denotes the weight link between the j-th hidden node and the k-th output node; $\theta_j^{(1)}$ and $\theta_k^{(2)}$ denote the thresholds of the j-th hidden node and the k-th output node. And φ and $\tilde{\varphi}$ denote the activation functions of hidden and output layers, respectively.

The activation functions of hidden and output layer are defined by the sigmoid function $\varphi(\gamma) = 1/(1+e^{-\gamma})$ and bi- tangent function $\tilde{\varphi}(\gamma) = \pi(1-e^{-\gamma})/(1+e^{-\gamma})$, respectively, where γ is the neuron input. The objective function of HPSONN is defined as the mean square error.

$$f = \frac{1}{P_t} \sum_{t=1}^{P_t} \sum_{k=1}^{n_{out}} (y_k^{\tau}(t) - y_k(t))^2 \tag{8}$$

where y_k^{τ} and y_k are the k-th desired and actual output of the neural network, respectively, P_t is the number of training set or test set.

4 Simulation and Results

Manipulator inverse kinematics control is realized by 3-layer FFNN and learned by using the hybrid particle swarm optimization. Fig.1. shows the model of the two links manipulator. Its forward kinematics equations can be represented by

$$\begin{aligned} x &= l_1 \cos q_1 + l_2 \cos(q_1 + q_2) \\ y &= l_1 \sin q_1 + l_2 \sin(q_1 + q_2) \end{aligned} \tag{9}$$

where l_1 and l_2 denote the length of the first link and the second link, respectively. And the coordinate x and y denote the position of the end effector.

The system of manipulator inverse kinematics control has two steps. First, in the learning step, the neural network can be learned by using the training joint angles and the positions of the end effector. Next, in the control step, the desired positions of the end effector are inputted to the trained neural network and it can generate the joint angles correspond to the desired positions.

The inputs of 3-layer FFNN are the manipulator position of *x* and *y* and the outputs are the joint angles. The training set and the test set are chosen as Ref. [17] and the number P_t of the training set and the test set are 36 and 900, respectively. The maximum evolution iteration is set 3000. The parameters C_1, C_2 are set 2.05 and 2.05, respectively. So φ is 4.1, *K* is 0.73 and the selection rate is 0.4. The optimization process of FFNN is shown in Fig.2. The manipulator position controlled by HPSONN and its actual position are shown in Fig.3.

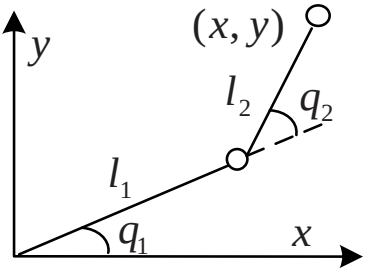


Fig. 1. The model of two links manipulator

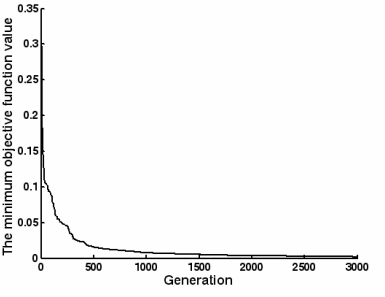


Fig. 2. The optimization process of NN

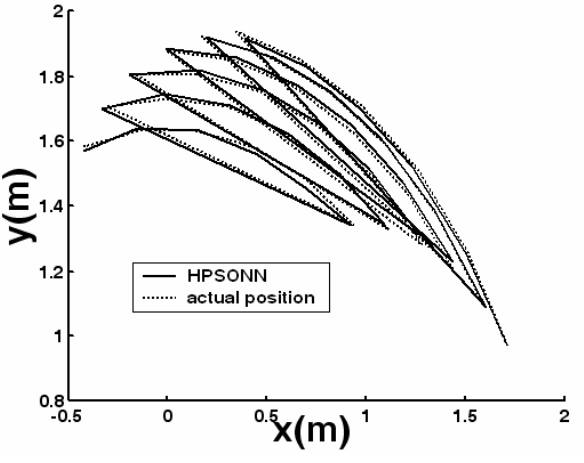


Fig. 3. Actual and HPSONN control position of manipulator

For comparison, we also use FBP, EGA, IGA and IEC optimize above FFNN. The training set and test set are the same as HPSONN. The control parameters of FBP, EGA, IGA and IEC are set as follows:

FBP: The maximum training iteration =20000, the square sum of permitted maximum error $\varepsilon = 0.01$, the learning rate $\eta = 0.01$, the increasing ratio of learning rate=1.03,

the decreasing ratio of learning rate=0.8, the moment constant=0.95 and the maximum error ratio=1.04. FBP algorithm ended in reaching the maximum training generation and the error was larger than the permitted maximum error.

EGA: The population size=80, the crossover probability=0.93, the mutation probability=0.08 and the maximum evolution iteration =10000.

IGA [18]: The population size =40, the offspring population size =40, the maximum evolution iteration =3000, the parameter $\alpha = 0.5$.

IEC[8]: The antibody size and clone size are set 10 and 40 respectively, suppression threshold ε_s = 0.02, the parameter α = 1. And the initial weight links are generated randomly in [-1, 1] ranges in above all methods.

Considering HPSO is also a stochastic optimization method and it is important to evaluate the average performance. We performed 10 trials in prescribed maximum iteration with initial position and initial velocity of all particles randomly generated in specified initial range, and the mean square errors of training sets and test sets of different methods are tabulated in Table 1. As seen in Table 1, the mean square errors of training sets and test sets of HPSONN for manipulator inverse kinematics control are much less than that of FBP, EGA and IGA, and it is slightly less than that of IEC. But HPSO takes two- thirds time of IEC to finish 3000 iteration. Therefore, HPSO has better generalization ability and can finish higher precision control of the manipulator.

However, the optimization result of HPSO is occasionally influenced by the parmeters of K and the selection rate and their values are set by experience. How to set the best value theoretically is the future work.

Table. 1. Results of different methods

Methods	Mean square error	
	Training sets	Test sets
HPSONN	1.23e-03	1.02e-03
FBP	0.1056	0.1272
EGA	8.68e-02	7.55e-02
IGA	8.71-03	7.96e-03
IEC	1.58e-03	1.26e-03

5 Conclusion

A hybrid particle swarm optimization based on the behaviour of insect swarms and natural selection mechanism is presented to optimize FFNN for manipulator inverse kinematics control. It provides a new parallel distributed model for computing the pseudoinverse Jacobian of redundant robot manipulators. By comparing the HPSO method with FBP, EGA, IGA and IEC, the mean square error of training sets and test sets are all smaller than above methods. It is proven that HPSO has better generalization ability and can finish higher precision control of the manipulator. It can also be used for other related optimial problems.

Acknowledgements. The research was supported by QingLan Project of Jiangsu Province, Key Research Project of Nanjing Institute of Technology(KXJ07078) and Remote Measurement and Control Key Lab Project of Jiangsu Province.

References

1. Tabandeh, S., Clark, C., Melek, W.: Genetic Algorithm Approach to Solve for Multiple Solutions of Inverse Kinematics using Adaptive Niching and Clustering. IEEE Congress on Evolutionary Computation, 1815–1822 (2006)
2. Poon, J.K., Lawrence, P.D.: Manipulator Inverse Kinematics based on Joint Functions. In: IEEE Int. Conf. on Robotics and Automation, vol. 2, pp. 669–674 (1988)
3. Brown, M., Harris, C.: Neuralfuzzy Adaptive Modeling and control. Prentice Hall, Englewood Cliffs (1994)
4. Hussein, S.B., Jamaluddin, H., Mailah, M.: A Hybrid Intelligent Active Force Controller for Robot Arms Using Evolutionary Neural Networks. In: Proc. of Congress on Evolutionary Computation, vol. 1, pp. 117–124 (2000)
5. Kim, J.H., Lee, C.H.: Evolutionary Ordered Neural Network and Its Application to Robot Manipulator Control. In: IEEE International Conference on Industrial Electronics, Control, and Instrumentation, vol. 2, pp. 876–880 (1996)
6. Hertz, J., Krogh, A., Palmer, R.G.: Introduction to the Theory of Neural Computation. Addison- Wesley, Reading (1991)
7. Li, M., Tam, H.Y.: Hybrid Evolutionary Search Method Based on Clusters. IEEE Trans. Pattern Analysis and Machine Intelligence 23, 786–799 (2001)
8. Wen, X.L., Xue, X.L., Zhang, P.: The Combination of Immune Evolution and Neural Network for Nonlinear Time Series Forecasting. In: Proc. of, Sixth International Symposium on Instrumentation and Control Technology, pp. 876–880 (2006)
9. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: Proc. IEEE Int. Conf. Neural Networks, vol. 4, pp. 1942–1948 (1995)
10. Seo, J.H., IM, C.H., Heo, C.C.: Multimodal Function Optimization Based on Particle Swarm Optimization. IEEE Trans. On Magnetics 42, 1095–1098 (2006)
11. Ren, P., Gao, L.Q., Li, N.: Transmission Network Optimal Planning Using the Particle Swarm Optimization Method. In: Proc. Int. Conf. on Machine Learning and Cybernetics, pp. 4006–4011 (2005)
12. Angeline, P.: Using Selection to Improve Particle Swarm Optimization. In: IEEE Int. Conf. Evol. Computation, Anchorage, Ak (1998)
13. Naka, S., Genji, T., Yura, T.: A Hybrid Particle Swarm Optimization for Distribution State Estimation. IEEE Trans. Power Syst. 18, 60–68 (2003)
14. Naka, S., Genji, T., Yura, T.: Swarm Intelligence. Morgan Kaufmann, San Mateo (2001)
15. Eberhart, R., Shi, Y.: Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization. In: Proc. Cong. Evolutionary Computation, pp. 84–88 (2000)
16. Baillieu1, J.: Kinematic Programming Alternatives for Redundant Manipulators. In: IEEE Int Conf. on Robotics and Automation, vol. 2, pp. 722–728 (1985)
17. Watanabe, B., Shimizu, H.: A Study of Generalization Ability of Neural Network for Manipulator Inverse Kinematics. In: IEEE Int Conf. on Industrial Electronics, Control and Instrumentation, vol. 2, pp. 957–962 (1991)
18. Wen, X.L., Song, A.G.: An Improved Genetic Algorithm for Plannar Straightness and Spatial Straightness Error Evaluation. International Journal of Machine Tools and Manufacture 43, 1077–1084 (2003)

Methods for Decreasing Time and Effort during Development and Maintenance of Intellectual Software User Interfaces

Valeriya Gribova

Institute of Automation and Control Processes
Far Eastern Branch of Russian Academy of Sciences,
Vladivostok, Russia
gribova@iacp.dvo.ru
http://www.iacp.dvo.ru/is/staff.php?id=11

Abstract. Methods for automated development and maintenance of intellectual software user interfaces are proposed. They rest on an ontology-based approach and are intended for decreasing effort and time for user interface development and maintenance. The principal conception of the approach, user interface project components, and methods for automated development of the project as well as a comparison of the methods and their analogues are presented.

Keywords: ontology, interface project, automated generation.

1 Introduction

The growing trend in software development is towards increase of the intellectual software based on knowledge and ontologies. Generally, an intellectual system consists of a user interface, an application program, ontologies, and knowledge. Their development, implementation, and posterior modification (maintenance) require great efforts and much time. The main reasons are:

- user interfaces are based on knowledge;
- knowledge bases about dialog for input/output parameters, for user intellectual support are large in size and/or have a complicated task model;
- knowledge about dialog for input/output parameters , for user intellectual support is extensive and/or has a complicated task model;
- user interfaces must be adaptable to variations in user requirements and knowledge modification during the software life cycle;
- user interfaces must be developed based on usability principles and an appropriate style guide.

Present-day tools for user interface development (Interface Builders and Model-Based User Interface Development Environments), as mentioned in current literature, do not provide a way to considerably decrease effort and time. Acording to some reviews (for example, [1]) the user interface development takes 50% time required to software development and its source code aggregates 48% of the total code.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 792–799, 2008.
© Springer-Verlag Berlin Heidelberg 2008

In view of the aforesaid, the urgency of research on decreasing labor- and time-consuming operations during development and maintenance of intellectual software user interfaces is obvious.

To solve the problem, an ontology-based approach to user interface development, automatic generation and maintenance has been proposed [2, 3]. The aim of this article is to describe methods for decreasing time and effort during development and maintenance of intellectual software user interfaces.

2 Ontology-Based Approach to User Interfaces Development

The basic ideas of the ontology-based approach are:

- to separate development of the user interface and the application program;
- to divide an interface project into components in according to conceptual framework of different groups of interface developers;
- to form the interface project by domain-independent ontologies (using a structural or graphical editor) that describe features of every component of the;
- based on this project to automatically generate the user interface code to a programming language;
- to link the interface code with the application code (a programming language, an interaction type with application, local or distributed, is determined by the developer).

As a result development and maintenance of a user interface reduce to development and maintenance of its project, because interface implementation is automatic. Using structural or graphical editors controlled by ontologies simplify a problem of interface project development, because the developer doesn't need to learn programming or specification languages.

3 User Interface Project

A user interface project defines a set of dialogs with the user for input/output parameters, control of the application program, and providing context-sensitive help for users [2, 3]. The interface project consists of a domain project, a task project, a presentation project, an application program project, a scenario dialog project, and a mapping project.

The domain project (DP) describes a set of domain terms consisting of input/output parameters of an application program, control information, and intellectual support of users activities. This component is specified by the domain ontology and is expressed by the pair: DP=(G, σ), where G is an oriented graph of the DP, σ is a graph labeling. The oriented graph G=<Vertices, Edges, RootVertex>, where Vertices is a set of graph vertices, Edges is a set of graph edges, RootVertex is a root vertex.

The set of graph vertices $\text{Vertices}=\{\text{Vertex}_i\}_{i=1}^{vertexescount}$ consists of two subsets – a set of terminal vertices and a set of nonterminal vertices, $\text{Vertex}_i \in$ Nontermi-

nal_Vertices ∪ Terminal_Vertices, where Nonterminal_Vertices is a set of nonterminal vertices, Terminal_Vertices – a set of terminal vertices.

The set of graph edges $Edges = \{Edge_i\}_{i=0}^{arcscount}$, $Edge_i = <_{VertexFromi}, VertexTo_i >$, where $VertexFrom_i \in$ Vertices, $VertexTo_i \in$ Vertices. The root vertex RootVertex is nonterminal vertex, RootVertex∈ Nonterminal_Vertices. The graph labeling σ is transformation of the vertex set and the edge set to a name set Ω, where Ω = DomainName $\cup\{GroupName_i\}_{i=1}^{termgroupcount} \cup \{TermName_j\}_{j=0}^{termcount} \cup \{TermAttribute_i\}_{i=1}^{attributecount} \cup \{QualityValue_i\}_{i=2}^{valuecount} \cup\{((RealMin, RealMax), RealMeasure)_i\}_{i=0}^{floatcount} \cup\{((IntegerMin, IntegerMax), IntegerMeasure)_i\}_{i=0}^{integercount} \cup\{N_i\}_{i=0}^{stringscount} \cup\{$TermGroup, Term, Attribute, Joint, Disjoint, RealValue, IntegerValue, SrtingValue}. The name set Ω is defined by names of the domain ontology, and graph labeling is defined by the vertex type $VertexFrom_i$ and $VertexTo_i$.

The Task Project (TP). The propose of an application program is to solving user tasks. Tasks could be divided into subtasks. Subtasks are steps for solving the main task. There are some relations among tasks. They determine conditions and the order of task performance [4]. TP=(T, σ), where T is a task tree, σ is a labeling tree. The task tree T=<Vertices, Edges, RootVertex>. The tree labeling σ is a transformation of the vertex set to a name set Ω, where $\Omega = \{TaskName_i\}_{i=1}^{taskcount} \cup$ SetType. The name set Ω consists of task names and set names. SetType ={«choice», «interleaving», «enabling», «deactivation»}. Labeling σ is defined as RootVertex→ CommonTaskName, the root vertex transforms to the common task name. Every vertex $Vertex_i$ → $(Label1_i, Label2_i)$, where $Label1_i \in \{TaskName_i\}_{i=1}^{taskcount}$, $Label2_i \in$ SetType (semantics of sets is defined in [4]).

The presentation project describes the structure and parameters of WIMP-interfaces. This component is specified by the WIMP-presentation ontology and is defined as a set of windows: $Windows = \{Window_i\}_{i=1}^{windowcount}$, where $Window_i \in$ Controls | Controltype = ContainerWindow. According to the WIMP-presentation ontology every interface element $Control_i$ is defined by a type, sets of parameters, functions, and events. The type, functions and events of an interface element have been defined in the ontology so the presentation project describes values of interface parameters for the interface project. Other interface elements could also be parameters of a window $Window_i$. In this case $Param_Type_k = Control_m$, where $Control_m \in$ Controls.

The application program project describes a method of interaction between the user interface and the application program, and program interfaces to provide this interaction. This component is specified by the corresponding ontology and is defined as a set of $Interfaces = \{Interface_i\}_{i=1}^{interfacecount}$. Every element of the set consists of the program interface definition assigned by the application program, $Interface_i$ = < Inter-

facename$_i$, InteractionModel$_i$, Functions$_i$>, where Interfacename$_i$– a program interface name. InteractionModel$_i$ is an element from the set IModels={Local, Distributed}; a function set Functions describes program interfaces assigned by the application program.

The scenario dialog project describes a first window StartWindow, specifying the beginning of dialog with users, StartWindow=Window$_i$, where Window$_i$ ∈ Windows, and so a set of possible conditions of dialog States ={State$_i$} $^{statecount}_{i=1}$. A state State$_i$ consists of a definition of the event Event$_i$, the set of variables Variables$_i$ for definition of a instruction sequence called Instructions$_i$ and the instruction sequence Instructions$_i$=<Instruction$_{ij}$> $^{instructioncount}_{j=1}$.
The instructions can be:

- program interface calls of the application program;
- system function calls;
- compound elements (“if” and “cycle” constructions) consisting of a sequence of the mentioned above elements.

The scenario dialog project is specified by the scenario dialog ontology.

The mapping project describes a mapping set among different elements of the interface project: elements of the domain and the task projects are mapped on parameters of interface elements in the presentation project, parameters of interface elements on output parameters of the application program in the scenario dialog project, and so on.

4 Methods for Specifying the Presentation Project

Application of the tool based on the ontological approach has shown that design and maintenance of the presentation component in the user interface project remains difficult for complicated and complex interfaces due to high requirements to developers. They are to know usability principals, various standards of development, take into account users’ requirements, conditions and environment of using software, and so on. As a result, development of methods and tools for automation of the presentation component design in the user interface project is an urgent task.
Interface element parameters of the presentation project are specified by the domain and the task projects, that is for every domain term (groups of terms) and every task (tasks) the interface developer chooses its (their) presentation in the interface, a set of interface elements with specified properties.

The following problems arise in this case:

- Time-consuming development of the presentation project for large size of the domain or the task projects;
- Labour-consuming maintenance of the interface project (modification of the domain project requires modification of the presentation project);
- High-level professional requirements to developers (they must know usability principals, requirements of standards, users, a domain for information presentation, and so on).

To meet the main requirement to the interface project, namely, easy modification and maintenance, some methods for specifying the presentation project are suggested. The choice of one of them is determined by the developer and depends on features of components of the interface project (for example the size of the presentation or task project, the domain features specifying type of information, and so on).

The Method for Direct Specifying the Presentation Project. In this case values of interface element parameters are direct references either to domain project elements or task project elements, that is, for an interface element of the presentation project $Control_i \in Windows$, where Param_Type=String, Param_Value=Value, $Value \in \{TaskName_i\}_{i=1}^{taskcount} \cup DomainName \cup \{GroupName_i\}_{i=1}^{termgroupcount} \cup \{TermName_j\}_{j=0}^{termcount} \cup \{AttributeName_i\}_{i=1}^{attributecount} \cup \{QualityValue_i\}_{i=2}^{valuecount}$.
For example, parameter values "Text" of the list box elements could be references to attribute values "localization" (frontal region, parietal region, left temporal region, right temporal region, occipital region etc.) from the domain project. The number of list box elements is determined by the number of "localization" attribute values; the modification of the number of this attribute values results in the modification of the number of list box elements; the modification of values of this attribute themselves in the domain project results in the modification of the "Text" parameter.

This method for specifying the presentation project is convenient for the small size of the domain or the task projects and also if accurate presentation of domain terms in an interface is required (for example, if the user sets special requirements).

The Method for Regular Specifying the Presentation Project. In this case values of interface element parameters are references to components of the domain ontology (but not to components of the domain project): groups of terms, terms, values, and attributes. That is, for an interface element of the interface project $Control_i \rightarrow Element \mid Control_i \in Windows, Element \in \{TermGroups, Terms, Attributtes, QualiyValues\}\}_{i=1}^{controlcount}$. As a result, every term group (terms, values, attributes) included in the domain project is represented by an interface element the developer chooses. In this case the value of a string parameter(s) is (are) elements of the domain project.

For example, the parameter values "Elements" = $\{ListBoxElement_i\}_{i=1}^{elemcount}$ of list box elements correspond to all quality joint values from the domain ontology (Joint values are a subset of the value set, disjoint value is only value of the value set). It means that all quality joint values from the domain project are presented by the same interface elements. At the same time their quantity could be arbitrarily large and is determined by the domain. This method for specifying the presentation project is convenient for the large size domain or projects (hundreds and thousands of terms).

The Method for Fragmentary Specifying the Presentation Project. In this case the developer divides the presentation project on a set of fragments based on user requirements, functionality and semantics of the application program; each fragment corresponds its representation in the interface: $\{Fragment_Presentastion \rightarrow Fragment_Terms \mid Fragment_Presentastion \in \{Control_i\}_{i=1}^{controlcount}, Control_i \in Windows,$

$\text{Fragment_Terms} \subset G\}^{i=1}_{fraggmentcount}$.

The principle statements of this method are:

- every fragment of the domain project automatically corresponds to a set of possible representations.
- corresponding is based on usability principals.
- if a fragment has some possible representations according to usability principles, then there are two variants:
- the developer chooses the only representation from the set of possible representations;
- the representation is automatically chosen according to default rules.

The explanation with the protocol of execution is given to the developer. The protocol consists of the set of criteria used with the description of accepted and unaccepted criteria (the criteria is chosen according to usability principles).

A set of criteria must be expanded (because usability principles are grown and modified).

To expand a set of possible representations a presentation mapping ontology is suggested. In terms of the ontology rules of corresponding domain terms to their possible representations (according to usability principles) are determined. In this way the set of mappings (the knowledge base) is formed. A mapping is described as a pair of a mapping condition and a set of its possible representations. The condition could consist of a set of conditions, each of them determines parameter values. Conditions are verified by the domain ontology. Parameters of conditions are determined in terms of the domain ontology. For example, a condition could be the number of joint values of a term from the domain project and a method their selection (joint or disjoint). Based on various values of these conditions, a set of representations is chosen (a set of radio-buttons, check boxes, a list box, a combo box, etc). Different fragmentations of the domain project lead to different presentation projects.

This method for specifying the presentation project is convenient (as well as the previous method) for the large size domain or task projects (hundreds and thousands of terms).

5 Discussion of Results

At present the approach and methods described above have been implemented in a tool for development, automatic generation and maintenance of user interfaces based on the ontology-based approach. This tool has been used to implement and maintain interfaces in a number of application programs.

In discussing the results obtained it is reasonable to compare them with methods implemented in Model-based interface development environments (MB-IDE) for development and automatic generation of user interfaces [5-7]. The other class of specialized tools, Interface Builders, does not support development interfaces based on knowledge.

To develop a user interface by MB-IDE the developer has to learn some specification languages, division a user interface into components does not correspond to conceptual

framework of different groups of interface developers so effort and time for user interface development and maintenance are considerably increasing.

The method for direct specifying the presentation project is implemented in some MB-IDE where string parameter values of interface elements are domain terms or users' tasks. However, any modification of domain terms or users' tasks require that string parameter values of interface elements be modified also, and the interface be recompiled. This results in laborious interface maintenance. In the suggested method string parameter values of interface elements are references to domain terms or users' tasks, so modification of domain terms or users' tasks results in automatic modification of the string parameter values of interface elements, and no recompilation is necessary.

The method for regular specifying the presentation project is suggested for the first time. It allows the developer to considerably decrease development and maintenance effort. For example, the domain project of the System for Intellectual Support for Urologists [8] consists of 700 terms and groups of terms, and of about 5000 possible values. In this case the method for regular specifying the presentation project provides fast specification and automatic maintenance of the presentation project.

The method for fragmentary specifying the presentation project has been used in some MB-IDE. With the help of this method developers determine a window (a set of windows) and information for every window (from the domain and the task projects). Then a presentation for this information is chosen automatically and the developer only enhances the presentation. On the one hand, such an approach decreases the time of presentation project development; on the other hand, the main problem for developers is difficult maintenance of the project, because modification of terms and tasks calls for automatic representation of this information and all enhancements are lost. Attempts to save results of enhancements have been made in some MB-IDEs. Nevertheless, they failed when new terms or tasks were added or deleted. Using references solves this problem. Another drawback of this method in some MB-IDE is strict embedment of mappings (a set of presentations) in the tool. As a result, the developer is not aware of these mappings and they could not be expanded. According to the method suggested all mappings are available in the knowledge base and could be viewed and modified.

Acknowledgments. The research was supported by the Far Eastern Branch of Russian Academy of Science, the grant "An Expandable System for Quality Monitoring".

References

1. Myers, B., Rosson, M.: Survey on User Interface Programming. Tech. Rpt. CMU-CS-92-113, Carnegie-Mellon, School of Comp. Sci (February 1992), http://citeseer.ist.psu.edu/myers92survey.html
2. Gribova, V., Kleshchev, A.: Control of User Interface Development and Implementation Based on Ontologies. Control Sciences 2, 58–62 (2006)
3. Gribova, V., Kleshchev, A.: From an Ontology-oriented Approach Conception to User Interface Development. International Journal Information Theories & Applications 10(1), 87–94 (2003)

4. Gribova, V.: Automatic Generation of Context-sensitive Help Using a User Interface Project. In: Proc. of XIIIth Intern. Conf. “Knowledge-dialog-solution”, Varna, vol. 2, pp. 417–422 (2007)
5. Puerta, A.: A Model–based Interface Development Environment. IEEE Software 14(1), 41–47 (1997)
6. Puerta, A.R.: Issues in Automatic Generation of User Interfaces in Model-Based Systems. In: Computer-Aided Design of User Interfaces, pp. 323–325. Universitaires de Namur, Namur (1996)
7. Szekely, P.: Retrospective and Challenges for Model-Based Interface (1996), http://citeseer.nj.nec.com/szekely96retrospective.html
8. Gribova, V., Tarasov, A., Chernyakhovskaya, M.: A System for Intellectual Support of Patient Examination Controlled by the Software & Systems, vol. 2, pp. 49–51 (2007)

Adaptive Hybrid SMC-SVM Control for a Class of Nonaffine Nonlinear Systems

Heng Li, Jinhua Wu, and Youan Zhang

Department of Control Engineering, NAAU
Yantai, 264001, China
pla.henrylee@gmail.com, {hywjhua,zya63}@sina.com

Abstract. An improved hybrid adaptive controller is proposed for a class of nonaffine nonlinear systems based on sliding mode control (SMC) and support vector machine (SVM) in this paper. The sliding mode controller ensures the robustness of the closed loop system while the support vector machine is used to adaptive tracking. Moreover, the Lyapunov stability theory is employed to grantee the global stability. The simulation results have shown the effectiveness of the proposed method.

Keywords: nonaffine nonlinear systems, sliding mode control, support vector machine control, hybrid adaptive control.

1 Introduction

For recent years, many research studies have focused on the use of neural networks in adaptive nonlinear control[1]-[8]. However, there are some weak points when using neural networks, such as difficulty of choosing the number of hidden units, overfitting, and local minima solutions.

Support vector machine is a technique based on statistical learning theory and structural risk minimization [9][10], which has been very successful in pattern recognition and function estimation problems. But there has been little research on the use of least square support vector machine in nonlinear adaptive control, especially for nonaffine nonlinear systems.

In this paper, an adaptive control design procedure based on support vector machine is presented for a class of nonaffine nonlinear systems. During the design, the implicit function theorem is used to attain the desired performances and the robustness of the closed loop system is guaranteed by a sliding mode control. To improve the approximation accuracy, the adaptation laws are deduced from the stability analysis in the sense of Lyapunov. The simulation results demonstrate the good tracking performances.

This paper is organized as follows: Section 2 provides the introduction of SVM for nonlinear function approximation. In section 3, the problem formulation and an adaptive controller combined SVM with SMC for nonaffine nonlinear system are presented. Simulation results performed on an illustrative example are demonstrated to show the effectiveness of the approach in section 4. Section 5 contains the conclusion.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 800–807, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2 SVM for Nonlinear Function Approximation

Given a training dataset D of l sample $\{(x_1, y_1), \cdots, (x_l, y_l)\}, x_i \in R^d, y_i \in R$, consider estimating the regression function in the following function set:

$$f(x) = w^T \phi(x) + \beta \tag{1}$$

where the nonlinear mapping $\phi : R^d \to R^h$ maps $x_i \in R^d$ into a high dimensional feature space, $w \in R^h, b \in R$.

The method of support vector machine is to minimize the empirical risk under the structure element defined by $w^T w \le C_n$ .In the case of least squares SVM, the optimization problem is defined as

$$\min_{w,b,\varepsilon} J_{LS}(w, \varepsilon) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{l} \varepsilon_i^2 \tag{2}$$

subject to the equality constraints:

$$y_i = w^T \phi(x_i) + \beta + \varepsilon_i \qquad i = 1, \cdots, l \tag{3}$$

where $\varepsilon = [\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_l]^T \in R^{l \times 1}$ denotes the error vector, γ denotes an arbitrary positive real constant.

The conditions for optimality lead to the solutions of the following KKT system[9]:

$$\begin{bmatrix} 0 & \vec{1}^T \\ \vec{1} & K + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{4}$$

where $y = (y_1, \cdots, y_l)^T$, $\vec{1}_{l \times 1} = (1, \cdots, 1)^T$, $\alpha = (\alpha_1, \cdots \alpha_l)^T$, and

$$K(x_i, x_j) = \phi^T(x_i) \phi(x_j), \quad i, j = 1, \cdots, l \tag{5}$$

which satisfies the Mercer condition.

The resulting LS-SVM model for function estimation becomes:

$$f(x) = \sum_{i=1}^{l} \alpha_i K(x_i, x) + \beta \tag{6}$$

where α_i, $i = 1, 2, \cdots, l$ and β are the solutions to the linear system (4).

As one of the most popular kernel functions in machine learning, Gaussian kernel function is selected for controller design in this paper. It takes the following form:

$$K(x_i, x_j) = \exp\left\{-\frac{\left\|x_i - x_j\right\|^2}{2\delta^2}\right\} \tag{7}$$

where δ denotes the kernel parameter.

3 SVM Control of Nonaffine Nonlinear Systems

3.1 Problem Statement

Consider a single-input single-output (SISO) nonlinear system described by the following differential equation[11]:

$$y^{(n)} = f(y, y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}, u) \tag{8}$$

where $y \in R$ is the measured output, $u \in R$ is the control input, $y^{(i)}, i = 1, \cdots, n$, is the i^{th} derivative of y, $f(\cdot)$ is an unknown smooth continuous nonlinear function.

Let $x = [x_1, x_2, \cdots x_n]^T = [y, y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}]^T \in R^n$ be the state vector, we may represent system (8) in a state space model form:

$$\begin{cases} \dot{x}_i = x_{i+1}, \ i = 1, \cdots, n-1 \\ \dot{x}_n = b(x, u) \\ y = x_1 \end{cases} \tag{9}$$

where $b(x, u)$ is also an unknown smooth continuous nonlinear function.

The objective is to develop a control law using sliding model control to enable the output y to follow a desired trajectory y_d and ensure the robustness of the closed loop system. Without loss of generality, we assume the following:

Assumption 1: All the state variables $x = [x_1, x_2, \cdots x_n]^T = [y, y^{(1)}, y^{(2)}, \cdots, y^{(n-1)}]^T$ are available;

Assumption 2: The desired trajectory y_d is available, continuous and bounded;

Assumption 3: $b_u = \partial[b(x, u)] / \partial u > b_0 \neq 0$;

Assumption 4: There exists a smooth function $\varphi(\xi) > 0$ such that $\left|\dot{b}_u / b_u\right| \leq \varphi(\xi)$.

3.2 Design of Adaptive Control Law

Define the desired trajectory vector x_d, the tracking error vector e and the switching surface s as:

$$x_d = [y_d, \dot{y}_d, \cdots, y_d^{(n-1)}]^T, e = x - x_d = [e_1, e_2, \cdots, e_n]^T, s = [\Lambda \quad 1]^T e \tag{10}$$

where $\Lambda = [\lambda_1, \lambda_1, \cdots, \lambda_{n-1}]^T$ is an appropriately chosen coefficient vector such that $e \to 0$ as $s \to 0$. Then the time derivative of the switching surface s can be written as:

$$\dot{s} = b(x,u) - y_d^{(n)} + [0 \quad \Lambda^T]e \tag{11}$$

Lemma 1[1]: For the system (1) satisfying A1 through A4, there exists a compact set Φ_u and an unique ideal input u^* such that, for all $x(0) \in \Phi_u$, the equation (11) can be expressed as the following form:

$$\dot{s} = -\varphi(x)s \tag{12}$$

which allows to obtain $\lim_{t\to\infty} |y(t) - y_d(t)| = 0$.

The previous lemma guarantees the existence of a control law guaranteeing the convergence of the tracking error toward zero. According to Section 2, we use a support vector machine to approximate the ideal law u^* as follow:

$$u^* = \sum_{i=1}^{l} \alpha_i^* K(x_i, x) + \beta + \varepsilon \tag{13}$$

where α^* is the optimal value of α and ε is the approximation error.

To guarantee the stability of the closed loop system, we add a supplementary signal u_{smc}. Hence, the proposed control law is given by:

$$u = u_{svm} + u_{smc} \tag{14}$$

$$u_{svm} = \sum_{i=1}^{l} \alpha_i K(x_i, x) + \beta \tag{15}$$

$$u_{smc} = -k_0 s - k_1 sign(s) \tag{16}$$

where, $k_0, k_1 > 0$ are design parameters.

Using the Mean Value Theory[11], there exists a positive constant $\lambda \in [0,1]$ such that

$$b(x,u) = b(x,u^*) + b_{u_\lambda}(u - u^*) \tag{17}$$

where $b_{u_\lambda} = \left.\frac{\partial[b(x,u)]}{\partial u}\right|_{u=u_\lambda}$ and $u_\lambda = \lambda u + (1-\lambda)u^*$.

According to the implicit function theorem [12], there exists a u^* such that

$$v + b(x, u^*) \equiv 0 \tag{18}$$

where

$$v = \varphi(x)s - y_d^{(n)} - [0 \quad \Lambda^T]e \tag{19}$$

So, Eq. (17) can be written as

$$b(x,u) = -v + b_{u_\lambda}(u - u^*) = -\varphi(x)s + y_d^{(n)} + [0 \quad \Lambda^T]e \tag{20}$$

From Eq. (12) and Eq. (20), we can obtain:

$$\dot{s} = -\varphi(x)s + b_{u_\lambda}(u - u^*) \tag{21}$$

Using Eq.(14)-(16), Eq. (21) becomes:

$$\begin{aligned} \dot{s} &= -\varphi(x)s + b_{u_\lambda}u_{svm} + b_{u_\lambda}u_{smc} - b_{u_\lambda}u^* \\ &= -\varphi(x)s + b_{u_\lambda}\sum_{i=1}^{l}\tilde{\alpha}_i K(x_i, x) + b_{u_\lambda}[-k_0 s - k_1 sign(s)] - b_{u_\lambda}\varepsilon \end{aligned} \tag{22}$$

which gives

$$b_{u_\lambda}^{-1}\dot{s} = -b_{u_\lambda}^{-1}\varphi(x)s + \sum_{i=1}^{l}\tilde{\alpha}_i K(x_i, x) - k_0 s - k_1 sign(s) - \varepsilon \tag{23}$$

Consider the following Lyapunov function:

$$V = \frac{1}{2}(b_{u_\lambda}^{-1}s^2 + \sum_{i-1}^{l}\tilde{\alpha}_i^{\,2}) \tag{24}$$

Differentiating Eq. (24) along Eq.(21) yields and using the fact that $\dot{\tilde{\alpha}}_i = \dot{\alpha}_i - \dot{\alpha}_i^* = \dot{\alpha}_i$, We have

$$\begin{aligned} \dot{V} &= sb_{u_\lambda}^{-1}\dot{s} + \sum_{i-1}^{l}\tilde{\alpha}_i\dot{\alpha}_i \\ &= s[-b_{u_\lambda}^{-1}\beta(x)s + \sum_{i=1}^{l}\tilde{\alpha}_i K(x_i, x) - k_0 s - k_1 sign(s) - \varepsilon] + \sum_{i-1}^{l}\tilde{\alpha}_i\dot{\alpha}_i \\ &= -b_{u_\lambda}^{-1}\beta(x)s^2 + s\sum_{i=1}^{l}\tilde{\alpha}_i K(x_i, x) - k_0 s^2 - k_1|s| - \varepsilon\, s + \sum_{i-1}^{l}\tilde{\alpha}_i\dot{\alpha}_i \\ &\le -k_0 s^2 - k_1|s| - \varepsilon\, s + \sum_{i=1}^{l}\tilde{\alpha}_i[\dot{\alpha}_i + sK(x_i, x)] \end{aligned} \tag{25}$$

Choosing the following adaptation law

$$\dot{\alpha}_i = -sK(x_i, x) \tag{26}$$

gives

$$\dot{V} \le -k_0 s^2 - k_1 |s| - \varepsilon\, s = -\frac{k_0}{2} s^2 - k_1 |s| - \frac{k_0}{2}(s - \frac{\varepsilon}{k_0})^2 + \frac{\varepsilon^2}{2k_0} \le -\frac{k_0}{2} s^2 + \frac{\varepsilon^2}{2k_0} \tag{27}$$

Integrating the above inequality, and noting the fact that $V(\infty) \ge 0$, we have

$$\int_0^\infty \frac{k_0}{2} s^2 dt \le V(0) - V(\infty) + \int_0^\infty \frac{\varepsilon^2}{2k_0} dt \le V(0) + \int_0^\infty \frac{\varepsilon^2}{2k_0} dt < \infty \tag{28}$$

which guarantees that $s \in L_2$.

According to Eq.(23), we have $\dot{s} \in L_\infty$, then using Barbalat Lemma, we have $s \to 0$ when $t \to \infty$. Therefore, the tracking error converge to the origin, i.e., $\lim_{t\to\infty} e = 0$.

Now, we get the following theorem:

Theorem 1: For the system (8) satisfying A1 through A4, if the control law is Eq. (14) and the adaptation law is Eq. (26), then, there exists a compact set Φ_u such that, for all $x(0) \in \Phi_u$, the tracking error can converge to the origin.

4 An Illustration Example

To illustrate the effectiveness of the proposed adaptive controllers, we consider the following system:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_1^2 + 0.1u^3 + 0.1(1 + x_2^2)u + \sin(0.1u) \\ y = x_1 \end{cases} \tag{29}$$

In this simulation example, the control objective is to determine u so that the output y follows the desired reference trajectory $y_d = \sin(2\pi t / 3)$. The plant must attain the following surface $s = \dot{e} + e$, and slide on it to reach the origin $\dot{e} = e = 0$. Fig.1 and Fig.2 present the numerical simulation results for the initial conditions $[x_1, x_2]^T = [0.5, 0.5]^T$. Fig. 3 gives the control signal. The SMC control parameters are $k_0 = 15$, $k_1 = 10$ and the SVM parameters are $l = 100$, $\delta = 0.6$, $\gamma = 1000$ and $\varepsilon = 0.001$. In addition, for the purpose to eliminate the chatter of SMC, we introduce

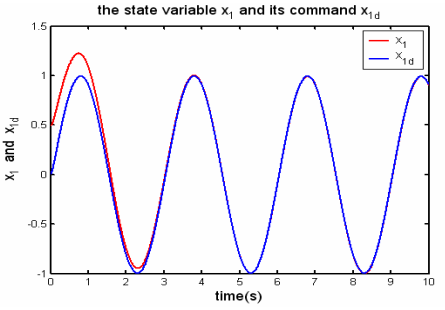


Fig. 1. Simulation result of variable x_1 and x_{1d}

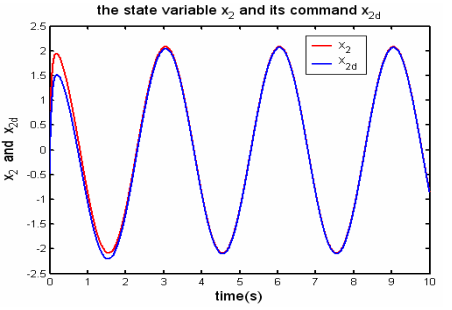


Fig. 2. Simulation result of variable x_2 and x_{2d}

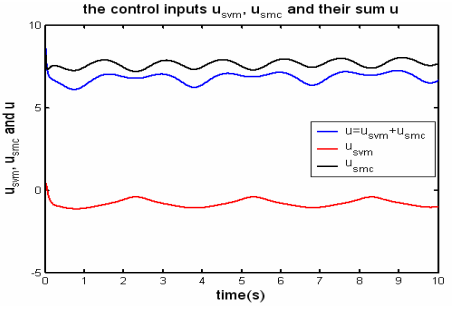


Fig. 3. Simulation result of control u_{SVM}, u_{SMC} and u

the function $\frac{s}{|s|+\eta}$, $\eta = 0.01$ to continue the function $sign(s)$. The simulation results are shown in Fig.1-Fig.3

Fig.1 and Fig.2 are the simulation results of variable x_1, x_2 and their desires x_{1d}, x_{2d} respectively and they show the good tracking performance and the convergence of the state variables to their reference trajectories. Fig.3 is the simulation results of control efforts u_{svm}, u_{smc} and the total control effort u, and it has clearly shown the less control effort by hybrid SMC-SVM than only by SMC.

5 Conclusions

In this paper an improved hybrid adaptive controller is proposed for a class of nonaffine nonlinear systems based on sliding mode control and support vector machine. The sliding mode controller ensures the robustness of the closed loop system while the support vector machine is used to adaptive tracking. Moreover, the Lyapunov stability theory is employed to grantee the global stability. The simulation results have shown the effectiveness of the proposed method. But the propose method uses the offline support vector machine, what should be researched in the future is using the on-line support vector machine to ensure the real-time performance.

References

1. Chen, F.C., Khalil, H.K.: Adaptive Control of Nonlinear Systems Using Neural Networks. International Journal of Control 55(6), 1299–1317 (1992)
2. Zhao, T., Sui, S.L.: Adaptive Control for A Class of Non-affine Nonlinear Systems via Two-layer Neural Networks. In: Proceedings of the 6th World Congress on Intelligent Control and Automation, Dalian, vol. 1, pp. 958–962 (2006)
3. Polycarpou, M.M.: Stable Adaptive Neural Control Scheme for Nonlinear Systems. IEEE Transaction on Automatic Control 41(3), 447–451 (1996)
4. Chen, H.-J., Chen, R.S.: Adaptive Control of Non-affine Nonlinear Systems Using Radial Basis Function Neural Network. In: Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control, Taipei (2004)
5. Ge, S.S., Wang, C.: Direct Adaptive NN Control of A Nonlinear Systems. IEEE Transaction on Neural Networks 13(1), 214–221 (2002)
6. Karimi, B., Menhaj, M.B., Saboori, I.: Robust Adaptive Control of Nonaffine Nonlinear Systems Using Radial Basis Function Neural Networks. In: Proceedings of the 32nd Annual Conference on IEEE Industrial Electronics (2006)
7. Guo, X.J., Han, Y.L., Hu, Y.A., Zhang, Y.A.: CMAC NN-based Adaptive Control of Non-affine Nonlinear Systems. In: Proceedings of the 3rd World Congress on Intelligent Control and Automation, Hefei, vol. 5, pp. 3303–3305 (2000)
8. Zhang, T., Ge, S.S., Hang, C.C.: Direct Adaptive Control of Non-affine Nonlinear Systems Using Multilayer Neural Network. In: American Control Conference, Philadelphia, vol. 1, pp. 515–519 (1998)
9. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (1995)
10. Suykens, J.-A.K., Vandewalle, J.: Least Square Support Vector Machine Classifiers. Neural Processing Letters 9, 293–300 (1999)
11. Khalil, H.K.: Nonlinear Systems, 3rd edn. Prentice-Hall, Englewood Cliffs (2002)
12. Goh, C.J.: Model Reference Control of Nonlinear Systems via Implicit Function Emulation. International Journal of Control 60(1), 91–115 (1994)

Hormone-Inspired Cooperative Control for Multiple UAVs Wide Area Search

Hui Peng, Yuan Li, Lin Wang, and Lincheng Shen

School of Mechanical Engineering and Automation,
National University of Defense Technology, China
penghui.ph@gmail.com

Abstract. Teams of unmanned aerial vehicles (UAVs) are well suited to perform a search mission where multiple unknown targets located over a wide area. It is an important research area in cooperative control. Hormones are the common biological mechanism that enables cells to communicate and coordinate without central units. In this paper, we combine digital hormone mechanisms with recent search map method to provide a more effective way for multiple UAVs cooperative search. By using this extended map, single UAV can make local path decisions and coordinate their search behavior with neighboring UAVs via digital hormones information. UAV teams can display a better area searching ability without central command and control. Simulation results demonstrate the efficiency of the proposed approach.

Keywords: UAV, cooperative control, wide area search, digital hormone.

1 Introduction

Cooperative Control of multiple unmanned aerial vehicles (UAVs) to performing complexity tasks has captured the attention of control community in recent years [1, 2]. Because of limited inter-vehicle sensing and communication, single UAV has only limited information, which leads to the interaction between UAVs occur locally. Wide area search is a typical cooperative control problem, where UAV teams act cooperatively to search for unknown targets in uncertain and dynamic environments. Each vehicle must plan a local path over the environment to meet this global goal. One of the challenges in achieving multiple UAVs cooperative search is to design coordination method so that local decision will result in group cooperation.

A great of works focuses on the problem of multiple UAVs wide area search. In [3], a general framework for conducting multi-agent cooperative search is proposed. In [4], a search-theoretic approach is taken. An agent based negotiation as the decision making mechanism to obtain search path is presented in [5]. Search map is also a useful method for UAV area search in many literatures [3, 6, 7, 8].

In recent years, self-organization (SO) is a promising method to this problem. Usually in some biological systems, such as flocks of birds, schools of fishes and ants swarm, each individual has very limited sensing and communication abilities. However, the system as a whole can display highly complex and collective behavior

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 808–816, 2008.
© Springer-Verlag Berlin Heidelberg 2008

without central control. Inspired by this collective behavior of biological systems, Parunak [9] used digital pheromone mechanisms as an extension of potential fields for UAV control, UAVs maintain a pheromone map to coordinate with each other and guide their individual behavior. Erignac [10] presented an exhaustive search strategy for distributed pheromone maps. Price [11] proposed a SO model based on GA algorithm, which allowed for the behaviors evolution of UAV swarms and enable the simulated UAVs to effectively search and destroy targets. Shen [12, 13] designed the Digital Hormone Model (DHM) for robot swarm and self-organization, the model allows many simple robots in large-scale systems to react to each other, and self-organize into global patterns that are suitable for task and environment.

In this paper, we extend the search map method by using Digital Hormone Mechanisms [12]. Then we present an optimal look-ahead search path planning algorithm based on extended search map. Our approach enables UAV to make its local search decisions and to coordinate with neighboring UAVs via hormone information more effectively.

The remainder of the paper is organized as follows. Section 2 presents the problem formulation for the UAVs and environment. Section 3 gives the detail of extended search map based on digital hormone. In Section 4, an ESM based path planning algorithm for UAV is proposed. Simulation results are presented and discussed in Section 5. Finally, Section 6 concludes the paper with a brief discussion of future direction for our work.

2 Problem Formulation

Consider a bound $L_x \times L_y$ area E in which N heterogeneous UAVs U_k, $k = 1,\ldots N$ to perform the cooperative search mission. Assume that there have M unknown targets T_m, $m = 1,\ldots M$ in the area. The goal of the UAV teams is to search and identify all targets in the environment with lower cost and less time. A desirable search operation for multiple UAVs would be to search the high uncertainty regions. The problem involves autonomous decision-making for on-line real-time path planning by each UAV in coordination with neighboring agents via communication.

2.1 Environment Description

Environment is usually represented in search map as a grid of cells, and each cell in the map associates with a certainty value $p_{ij} \in [0,1], i \in \{1,...,L_x\}, j \in \{1,...,L_y\}$. We suppose each cell contains at most one target, then $p_{ij}(t) = 0$ represents that the UAV knows nothing about the target in grid (i, j) at time step t, and $p_{ij}(t) = 1$ represents high probability that a target is present in grid (i, j). For every UAV, there exists such a map that stores search information about the environment. As the UAV moves around in the search area, it gathers new information about environment by Air-borne sensor and by communication with other UAV. Therefore, the search map is continuously evolving as new information is collected and processed. At time step t+1, the search map update is:

$$p_{ij}(t+1)=\begin{cases}\tau\cdot p_{ij}(t) & if\ \ no\,UAV\ visits\\ F(p_{ij}(t),z_{ij}) & otherwise\end{cases} \tag{1}$$

Where $\tau\in[0,1]$ is a dynamic information factors. If no UAV visits grid (i,j), the certainty value will decrease. If UAV visits grid (i,j), the certainty value depends on the UAV sensor observation $z_{ij}\in[0,1]$, and $z_{ij}=1$ indicates target detection. The sensor is assumed to be imperfect, with a parameter α characterizing its detection [6], the update function $F(\cdot)$ is:

$$F(p_{ij}(t),z_{ij})=\begin{cases}\dfrac{\alpha p_{ij}(t)}{1+(\alpha-1)p_{ij}(t)} & if\ \ z_{ij}=1\\ \dfrac{1-p_{ij}(t)}{1+(\alpha-1)p_{ij}(t)} & if\ \ z_{ij}=0\end{cases} \tag{2}$$

2.2 UAV Model

(1) *Dynamics Model*. We assume that the UAVs move on continuous trajectories with constant speed in the search area. For the k^{th} UAV U_k, a common motion model is given by the following equations [14]:

$$\begin{aligned}\dot{x}_k&=v_k\cos\theta_k\\ \dot{y}_k&=v_k\sin\theta_k\\ \dot{\theta}_k&=u_k\cdot\eta_{max}\quad |u_k|\le 1\\ \dot{v}_k&=0\end{aligned} \tag{3}$$

Where (x_k,y_k) is the position of U_k, also denoted by $p_k=(x_k,y_k)$; v_k is the speed, θ_k is its heading angle and u_k is the UAV decision variable, noticed that the turning rate must fulfill max value constraint η_{max}.

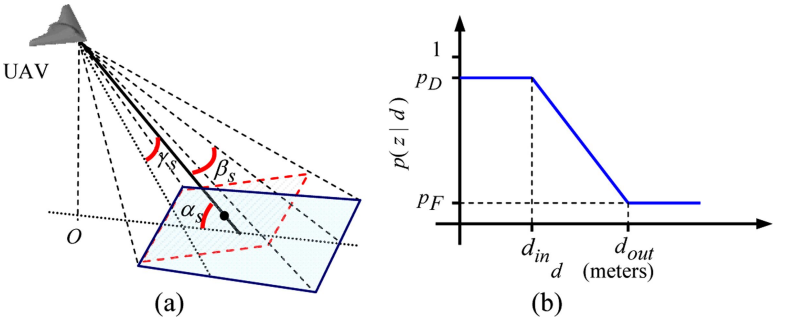


Fig. 1. Sensor Model

(2) *Sensor Model*. Each UAV is assumed to have a EO/IR system, the sensor has a limited field-of-view (FOV) and therefore it is only possible to see features that are inside the sensor footprint on the ground (as shown in Fig.1 (a)). In the sensor

footprint, UAV has some detection method which analyzes the incoming image and alarm if they find any target. Assume the sensor detection probability is $p(z \mid d)$. Where d is the distance between sensor and target, and $z \in [0,1]$ is the detection event, p_D is the sensor detection probability while target existence, and p_F is the false alarm probability. Then, the sensor parameter is defined as $\alpha = p_D / p_F$. The sensor probability model [15] is shown in Fig.1 (b). When the sensor is in its optimal range d_{in}, the probability of target detection is constant, if outside this best range, the target detection probability drops off quickly until a second threshold d_{out} is passed.

(3) *Communication*. Communication between UAVs is assumed to occur over a broadcast network that is noise free and instantaneous. However, communication is range-limited. The communication range of UAV U_k is r_k. We assume that all UAVs have the same communication range, $r_k = r, k = 1,...,N$, UAVs can communicate mutually within range of each other. Then, the neighboring node set for the vehicle U_k is $N_k = \{U_m \mid \sqrt{(x_k - x_m)^2 + (y_k - y_m)^2} < r_k, m = 1,2,...,N\}$.

3 Hormone Based Extended Search Map

Based on information of local search map, UAV can make its own path decisions. However, for multiple UAVs wide area search, cooperation between agents is a key issue. UAVs must have ability to coordinate their actions and to maximum team search efficiency. To achieve effective multiple UAV cooperation search, we extend search map based on digital hormone model.

3.1 Digital Hormone Model

In biological systems, cells are the basic component. Different cells secrete hormones and respond to different hormones, and large numbers of cells communicate with each other via reaction and diffusion of hormones. Inspired by hormone information, Shen [12, 13] proposed the Digital Hormone Model, which is a distributed control method for robot swarm. The model consists of three components: a specification of a dynamic network; a probabilistic function for individual robot behavior; and a set of equations for hormone reaction, diffusion and dissipation. Robots are viewed as biological cells that communicate and collaborate though hormones, and execute local actions via receptors. All robots have the same decision-making protocol, but they will react to hormones according to their local topology and state information so that hormone can cause robots swarm to perform different actions.

Mathematically speaking, DHM is in fact a kind of digital potential field method. The environment is represented and abstracted as a digital field via hormone mechanism. Single agent can make its own decisions through sensing the state of field, Because of positive feedback of hormone information, it lets the system (agents and environment) to auto-organize, so that the optimal coordination of multiple agents is finally achieved.

3.2 Extended Search Map

The extended search map (ESM) is a two-dimensional field $p'_{ij}(t) = [p_{ij}(t), H_{ij}(t)]$, where $H_{ij}(t)$ is the digital hormone information for grid (i, j). The concentration of hormone is a function of UAV position and of time. When UAV moves to (i, j), it generates digital hormone information H in its search map, and broadcasts hormone message to its neighbors. The information consists of two types of hormones, the activator H_A and the inhibitor H_I. The diffusion equation of H_A and H_I are given bellow [12]:

$$\begin{aligned} \Delta H_A(i, j, t) &= \frac{a_A}{2\pi\sigma^2} e^{-\frac{(i-a)^2+(j-b)^2}{2\sigma^2}} \\ \Delta H_I(i, j, t) &= -\frac{a_I}{2\pi\rho^2} e^{-\frac{(i-a)^2+(j-b)^2}{2\rho^2}} \end{aligned} \tag{4}$$

where grid (a,b) is adjacent to grid (i, j), a_A, a_I are constants. σ, ρ are the rates of diffusion respectively, and $\sigma < \rho$ to satisfy the Turing stability condition. With the diffusion of hormone in search map through communication, we get the hormone update function at time *t* by summing up all hormone information from neighboring UAV, where the constant $\tau_H \in [0,1]$ is the rate for dissipation.

$$H_{ij}(t+1) = \tau_H \cdot H_{ij}(t) + \sum_{N_k} (\Delta H_A(i, j, t) + \Delta H_I(i, j, t)) \tag{5}$$

4 Search Path Planning

The goal of search path planning is to generate the effective trajectory along with an objective function is maximized. To meet the requirement of UAV real-time searching, we use a limited look-ahead policy [3] to select next q time steps path, where the path planning is done in discrete time in a receding horizon manner.

4.1 Optimal Look-Ahead Search

Usually, the choices available to a UAV may be {turn left, go straight, turn right}, so the control decision of the k^{th} UAV is $u_k \in \{-1, 0, 1\}$. At each decision time step *t*, the path planner designs the path for steps *t*+1, *t*+2, …, *t*+*q*, where *q* is the planning horizon. Then, the control sequence for next q time steps is $\{u_k(t+1), \cdots, u_k(t+q)\}$. By considering the speed v_k and the maximum turning angle η_{max}, the new waypoint $\{p_k(t+1), \cdots, p_k(t+q)\}$ of the UAV at each time step within the planning horizon can be easily identified. Fig. 2 shows an example of the tree of waypoints for q = 3, it has 27 possible paths for UAV to choose.

Given the search tree at time step *t*, each UAV uses a depth-first optimization method to calculate the optimal local control sequence $\{u_k^*(t+1), \cdots, u_k^*(t+q)\}$, then it gets the optimal search path by executing the control sequence. Implementation of the UAV search path planning algorithm is described in Fig.3.

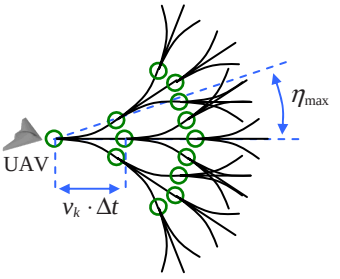


Fig. 2. Tree of Waypoints for the Path Decision

(1) Initialize the ESM $p'_{ij}(0)$, Set decision time $t = 0$
(2) **Repeat**
(3) Expand the path decision tree of next q step
(4) Calculate control sequence u_k^* using depth-first optimization
(5) **For** $i = 1$ to q **do**
(6) Execute the control input $u_k^*(t+i)$
(7) Update the extend search map $p'_{ij}(t+i)$
(8) **End for**
(9) $t = t + q$
(10) **Until** mission is finished

Fig. 3. ESM and optimal look-ahead search based path planning algorithm

4.2 Cost Function

The multi-objective cost function J can be written as the expected rewards associated with each of the possible paths at next time step. For the n^{th} possible path $\{p_k^n(t+1), \cdots, p_k^n(t+q)\}$, the cost function is given as:

$$J_n = \omega_1 J_{Tn} + \omega_2 J_{En} + \omega_3 J_{Cn} \tag{6}$$

where J_{Tn} is the target finding reward of the n^{th} path, J_{En} is the environment search reward, J_{Cn} is the cooperation reward, and $0 \le \omega_i < 1, i = 1,2,3$ is the corresponding weight. Based on the ESM information, the definitions of three rewards are given below, where R_n is the region covered by the n^{th} path.

(1) *Target Finding Reward*. To achieve the goal of maximizing the number of confirmed target, UAV will get reward if it can find a new target in one grid. The expected target finding reward is given as equation (7).

(2) *Environment Search Reward*. For area search purpose, it is better for the UAVs to visit grids with lower certainty. The actual certainty change is based on the

$$J_{Tn} = \sum_{(i,j)\in R_n} [(p_D - p_F)\, p_{ij}(t) + p_F] \tag{7}$$

sensor observation by equation (1) and (2). Here we use an estimate model for expected certainty change when a UAV will visit grid (i, j) at next time step. The estimate model is $p'_{ij}(t+1) = p_{ij}(t) + 0.5(1 - p_{ij}(t))$. Then the environment search reward is defined as the expected certainty increase:

$$J_{En} = \sum_{(i,j)\in R_n} [p'_{ij}(t+1) - p_{ij}(t)] = \sum_{(i,j)\in R_n} 0.5 \cdot (1 - p_{ij}(t)) \tag{8}$$

(3) *Cooperation Reward*. The cooperation between UAVs is to reduce the possible overlaps on their paths caused by their intentions to obtain high individual rewards. Digital hormone describes the relation of all UAVs' states. Repulsive hormones will push them apart when UAVs move close to each other. The balance between repulsive and attractive hormones ensures the cooperation of UAVs. The cooperation reward is defined as the function of hormone:

$$J_{Cn} = \sum_{i=2}^{q} [H(p_k^n(t+i)) - H(p_k^n(t+i-1))] \tag{9}$$

where $H(p_k^n(t))$ is the digital hormone value at the grid associated with waypoint $p_k^n(t)$ of the n^{th} path.

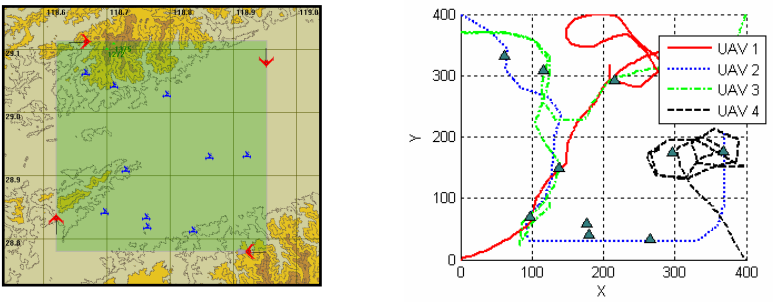
Noted that if $\omega_3 = 0$, the path planning algorithm will take no account of hormone information, and the ESM becomes the common search map; if the planning horizon $q = 1$, the limited look-ahead search becomes a greedy search in which the UAV moves at each step to a grid with highest reward.

5 Simulation Results

In this section, several simulation studies are performed to evaluate the efficacy of approach described above. We consider the following assumptions: 10 stationary targets are generated randomly within a 40km×40km area, the area is divided into 400×400 grids ; 4 UAVs are initially located at the four corners of the search region, and move at 120 m/s with $\eta_{max} = 70^\circ$. The communication range $r_k = 10$ km and sensor parameter $p_D = 0.8$, $p_F = 0.2$, $\alpha = 8$. The environment in which UAVs are searching is shown in Fig.4 (a), and Fig.4 (b) shows the total paths of 4 UAVs in one typical simulation, it can be seen that most of the targets are visited.

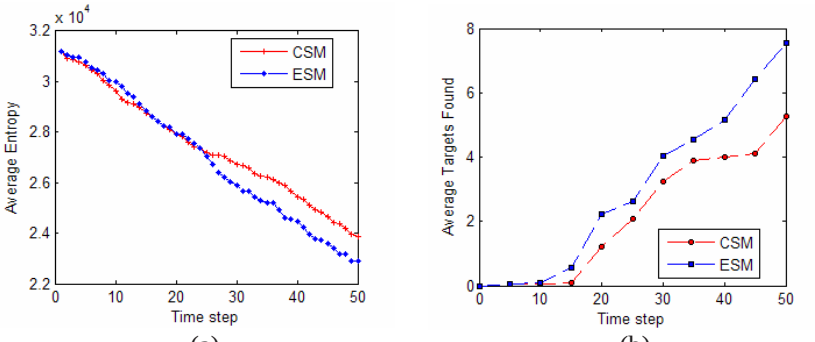
To illustrate the efficacy of our extended search map, we compare its performance with common search map method. The performance is measured from two aspects: the uncertainty of the search area and the number of targets found. The uncertainty can be defined as the information entropy of the map, which is given as:

$$e(t) = -\sum_{(i,j)\in E} (1 - p_{ij}(t)) \ln(1 - p_{ij}(t)) \tag{10}$$



(a) (b)

Fig. 4. The Simulation Environment



(a) (b)

Fig. 5. Simulation Results

The reducing of information entropy indicates the uncertainty decrease of environment. We run the simulation for 10 times under the same condition, where $\tau = 0.98$, $\tau_H = 0.9$, simulation time is 1500 seconds, decision time step is 30 seconds, and planning horizon $q = 3$. Fig.5 (a) shows the average decrease of entropy with increasing time for two map-building methods, and Fig.5 (b) shows the increment of average detected targets as a function of time. Form the simulation results, we can find that: Initially, as there is a lot of uncertainty in the area, two methods do well. However, as more of the environment is covered, the ESM based cooperative search becomes increasingly better at seeking out and covering regions of higher uncertainty, it enables the UAVs to detect more targets and to obtain more environment information. The results clearly show the superiority of ESM over common search map (CSM).

6 Conclusion

In this paper, an extended search map method for multiple UAVs cooperatively searching an unknown area is presented. The ESM combines the validity of certainty map with the coordination ability of hormone and provides an effective means for cooperative search problem. We illustrate how to utilize the extended search map to

UAVs path planning procedure, and present an ESM and optimal look-ahead search based path planning algorithm. The simulation results demonstrated that ESM is a promising and effective one. However, several issues remain to be addressed for multiple UAV cooperation. These include maintaining a consistent search map under imperfect network communication, and achieving consensus decision for distributed platforms. These problems are more difficult and practical for the applications of multiple UAVs and important directions for our research.

References

1. Chandler, P.R., Pachter, M., Swaroop, D., Fowler, J.M., Howlett, J.K., Rasmussen, S., Schumacher, C., Nygard, K.: Complexity in UAV cooperative control. In: American Control Conference, pp. 1831–1836 (2002)
2. Ryan, A., Zennaro, M., Howell, A., Sengupta, R., Hedrick, J.K.: An overview of emerging results in cooperative UAV control. In: IEEE Conference on Decision and Control, pp. 602–607 (2004)
3. Polycarpou, M.M., Yanli, Y., Passino, K.M.: A Cooperative Search Framework for Distributed Agents. In: IEEE International Symposium on Intelligent Control, pp. 1–6 (2001)
4. Baum, M.L., Passino, K.M.: A Search-Theoretic Approach to Cooperative Control for Uninhabited Air Vehicles. In: AIAA Guidance, Navigation, and Control Conference, pp. 1–8 (2002)
5. Sujit, P.B., Ghose, D.: Multiple UAV Search Using Agent Based Negotiation Scheme. In: American Control Conference, pp. 2995–3000 (2005)
6. Jin, Y., Minai, A.A., Polycarpou, M.M.: Cooperative Real-Time Search and Task Allocation in UAV Teams. In: IEEE Conference on Decision and Control, pp. 7–12 (2003)
7. Yanli, Y., Polycarpou, M.M., Minai, A.A.: Multi-UAV Cooperative Search Using an Opportunistic Learning Method. ASME Journal of Dynamic Systems, Measurement, and Control 129(5), 716–728 (2007)
8. Bertuccelli, L., How, J.P.: Search for Dynamic Targets with Uncertain Probability Maps. In: American Control Conference, pp. 737–742 (2006)
9. Parunak, H.V.D., Purcell, M., O'Connell, R.: Digital Pheromones for Autonomous Coordination of Swarming UAV's. In: AIAA Unmanned Aerospace Vehicles, Systems, Technologies, and Operations Conference, pp. 1–9 (2002)
10. Erignac, C.A.: An Exhaustive Swarming Search Strategy based on Distributed Pheromone Maps. In: AIAA Infotech@Aerospace Conference and Exhibit (2007)
11. Price, I.C.: Evolving Self-Organized Behavior for Homogeneous and Heterogeneous UAV or UCAV Swarms. M.S. Thesis. Air Force Institute of Technology (2006)
12. Wei-Min, S., Peter, W., Aram, G., Cheng-Ming, C.: Hormone-Inspired Self-Organization and Distributed Control of Robotic Swarms. Autonomous Robots 17(1), 93–105 (2004)
13. Feili, H., Wei-Min, S.: Mathematical foundation for hormone-inspired control for self-reconfigurable robotic systems. In: IEEE International Conference on Robotics and Automation, pp. 1477–1482 (2006)
14. Jin, Y., Liao, Y., Minai, A.A., Polycarpou, M.M.: Balancing Search and Target Response in Cooperative Unmanned Aerial Vehicle (UAV) Teams. IEEE Transactions on Systems, Man, and Cybernetics 36(3), 571–587 (2006)
15. Kamrani, F.: Using On-line Simulation in UAV Path Planning. M.S. Thesis. Sweden Royal Institute of Technology (2007)

A PID Parameters Tuning Algorithm Inspired by the Small World Phenomenon

Xiaohu Li[1], Haifeng Du[2], Jian Zhuang[1], and Sunan Wang[1]

1 School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi Province China
lxhxjtu@stu.xjtu.edu.cn
[2] Center for Administration and Complexity Science, Xi'an Jiaotong University, Xi'an 710049, Shaanxi Province, China
haifengdu@mail.xjtu.edu.cn

Abstract. Based on the small world phenomenon in social networks, a local search operator and a random long-range search operator with decimal-coding are proposed, then a decimal-coding small world optimization algorithm (DSWA) is explored, which validity and feasibility are testified by the simulation of PID parameters tuning. Compared with the corresponding genetic algorithm (GA) and particle swarm optimization algorithm (PSO), the simulation results indicate that DSWA can avoid trapping to local optimum in a certain extend, and it also can find high quality parameters for PID controller.

Keywords: PID Parameters Tuning, Small World Phenomenon, Decimal-Coding.

1 Introduction

The traditional PID controller is widely used in the field of industrial control, because of its simple structure, better robustness, facilitated tuning [1], etc.. As industrial systems are often nonlinear and uncertain, it is very difficult to apply traditional PID controller to achieve the desired results. So the optimization of tuning PID parameters is a very important research topic. Along with the development of control theory, combining these theories with traditional PID control strategy has come into being many new PID controllers, which make the performance of PID controllers improved greatly, all of them, intelligent control theory combining traditional PID controller is the most typical, which produced intelligent PID controller, for example: the fuzzy PID controller, the neural networks PID controller, etc.. However, the design of these intelligent PID controllers requires certain conditions which tend to restrict the application of them. Genetic algorithm (GA) which inspired by biological phenomena and particle swarm optimization algorithm (PSO) which suggested by physical phenomena have strong search capabilities, and have no special requirements on the optimization goal, when used to tuning PID parameters [2-3], they can achieve better results. In fact, Human social networks of which one of the famous examples is the small world phenomenon, may also suggest optimization mechanisms, which reveal that it can find the shortest path only relying on local connections and a few random long

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 817–824, 2008.
© Springer-Verlag Berlin Heidelberg 2008

connections in social networks [4]. Inspired by this, the literature [5] proposed an "imitation society" intelligent random search algorithm by simulating the small world phenomenon, which is named small world algorithm (SWA). When SWA is used to solve function optimization problems, the results show that SWA algorithm has some advantages which include the diversity of population, and fast convergence, and effectively overcoming the GA deceptive problem. However, the SWA is used in practical issues, such as the PID parameters tuning is still relatively rare.

In this paper, a decimal-coding SWA (DSWA) is designed on the basis of binary-coding SWA, which will be used for PID parameters tuning. The remainder of this paper is organized as follows: Firstly, Section 2 introduced the problem of PID parameters tuning. Then, in section 3, the process of PID parameters tuning with the small world optimization is presented, the local search operator and random long-range search operator are designed on the basis of the decimal-coding, then a decimal-coding small world optimization algorithm is proposed. By use of the error, control output, the rise time as an evaluation function for testing the performance of control system, two typical systems are treated as the objective systems for PID parameters tuning testing in section 4. Finally, section 5 is the conclusions of this paper.

2 Problem of PID Parameters Tuning

PID control strategy includes three parts: proportional, integral and differential, its basic form is given as the following equation:

$$u(t) = K_p\left[e(t) + \frac{1}{T_i}\int_0^t e(t)dt + T_d\frac{de(t)}{dt}\right] \tag{1}$$

Where, $u(t)$ is the output of the controller, $e(t)$ is the error of the system, proportional gain K_p, integral gain $K_i=K_p/T_i$, differential gain $K_d=K_pT_d$. When using PID strategy to control the system, it aims to make the system have better dynamic and static characteristics. Step signal is often used as the input, in order to obtain satisfied dynamic characteristics in the process of transition, the rise time t_r requires as small as possible to improve the response speed of the control system. The overshoot M_p must be small, a penalty function is adopted to avoid big overshoot. In addition, the controller output $u(t)$ must be limited in a certain scope in the actual system, and as small as possible to energy saving. As a compromise between these factors, the equation (2) is used as the evaluation criteria for the control system [6].

$$J = \begin{cases} \int_0^\infty \left(a_1|e(t)| + a_2u^2(t)\right)dt + a_3t_r & \text{if } e(t) \geq 0 \\ \int_0^\infty \left(a_1|e(t)| + a_2u^2(t) + a_4|e(t)|\right)dt + a_3t_r & \text{if } e(t) < 0 \end{cases} \tag{2}$$

Where, a_1, a_2, a_3, a_4 are the weight coefficients. In the event of overshoot, if $e(t)<0$, a penalty function will be added in the evaluation criteria. $a_4>>a_1$, this will help to restrain the overshoot firmly. Clearly, the objective of PID parameters tuning is an optimization problem which objective is to minimize J. Supposed the vector $\boldsymbol{x}=(K_p, K_i,$

$K_d) \in \Omega$ consists of the three PID parameters, which is in feasible solution set Ω. Because the value of the rise time, overshoot and controller output rely on the value of $\boldsymbol{x}$, The PID parameters tuning problem can be transformed into an optimization problem, which seeking the right vector $\boldsymbol{x}$ to minimize the objective function $J(\boldsymbol{x})$, $\overline{\boldsymbol{x}} \le \boldsymbol{x} \le \underline{\boldsymbol{x}}$, $\overline{\boldsymbol{x}} = (\overline{x}_1, \overline{x}_2, \overline{x}_3)$ and $\underline{\boldsymbol{x}} = (\underline{x}_1, \underline{x}_2, \underline{x}_3)$ are the lower and upper bounds of $\boldsymbol{x}$. Due to the complexity of the objective function, there are usually a number of local minimum which irregularly distributing, some traditional algorithms are prone to trap in local minimum, so they can not achieve a satisfied solution. However, the small world algorithm can keep good diversity, and has strong optimization capabilities for the problem with much extremum, therefore, the algorithm inspired by the small world phenomenon can be introduced for PID parameters tuning.

3 Decimal-Coding SWA for PID Parameters Tuning

3.1 PID Parameters Tuning with the Small World Optimization Process

The small world phenomenon is originated from the research on "tracking the shortest paths in American social networks" by a social-psychologist Stanley Milgram in 1960's [4]. It was found that on average any two persons in the world could be linked by six acquaintances, so it indicates that there are short paths in the social networks. Kleinberg pointed out that it could purely rely on local information for efficient search, and constructed a "two dimensional lattice small-world network" model [7]. These researches reveal the principle of small world phenomenon: in the complex social networks, through many local connections and a few random long connections, it will increase the efficiency of the information transmission. This paper simulates the principle of small world phenomenon to find short path for designing an algorithm on the basis of literature [5], and uses it to tune the PID parameters.

As the vector $\boldsymbol{x}$ with 3-dimensional PID control parameters affects the objective function $J(\boldsymbol{x})$, K_p, K_i and K_d can be quantized according to the precision of solutions, the feasible parameters solution set Ω would be spread into three-dimensional discrete grid community space, so the feasible solution set Ω are mapped as the feasible grid nodes set S in the community space. Therefore, after initializing the vector set of the candidate parameters $\Omega_1 \subseteq \Omega$, each candidate parameters vector $\boldsymbol{x} \in \Omega_1$ can be seen as a candidate node $\boldsymbol{s} \in S$ in the social grid space, the optimization process of PID parameters tuning becomes a candidate node (the candidate parameters vector) to transmit information to the optimal node (the optimal parameters vector) in networks. The whole process simulates small world phenomenon, which emphasizes local connection search, and supplements by a few random long connection search. Obviously, the search process of optimization objective function for PID parameters tuning follows by local search to the overall results, which is the principle of the small world phenomenon.

In Figure 1, each node has independent local search capability and certain global search capability, so it can realize a good parallel performance. It also can efficiently search in the entire feasible nodes set, because there is no selection operation in overall nodes for updating the new nodes, each node's information can be protected well,

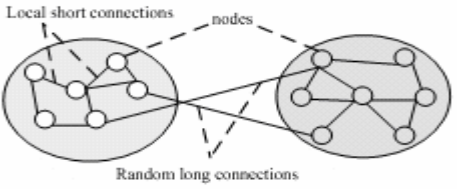


Fig. 1. The basic schematic of small world optimization process

it will no doubt increase the diversity of the candidate solution set, which help to avoid trapping in local optima.

3.2 The Strategy of Decimal-Coding

As PID parameters tuning is actually a multidimensional parameters optimization problem, in order to avoid using of binary-coding which will cause complexly in decoding and encoding, this paper adopted a decimal-coding strategy, that is, variable parameters coding strings use 0 to 9. After initializing the parameters as candidate solution set Ω_1, it can be mapped into the nodes set $S=\{\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_m\}$ in community grid networks, m is the number of nodes, the nodes $\boldsymbol{s}_i \in S$ are corresponding with the solution vector $\boldsymbol{x}_i \in \Omega_1$. When using the decimal-coding strategy, the code of the rth dimension parameter in node $\boldsymbol{s}_i$ is $s_r^i = s_1^i s_2^i \cdots s_j^i \cdots s_{l_i}^i$, $r \in \{1,2,3\}$, $s_j^i \in \{0,1,\cdots,9\}$, and the code length l_i is related to the precision of the solution. If the rth dimension parameter of vector $\boldsymbol{x}_i$ is limited in $x_r^i \in \left[\overline{x}_r^i, \underline{x}_r^i\right]$, the decimal-decoding equation is:

$$x_r^i = \underline{x}_r^i + (\overline{x}_r^i - \underline{x}_r^i)\sum_{j=1}^{l_i}(s_j^i \times 10^{-j}) \qquad i = 1,2,\cdots,m;\ r = 1,2,3. \tag{3}$$

3.3 Local Search Operator and Random Long-Range Search Operator

Because of the decimal-coding, $\mathrm{d}(\boldsymbol{s}_i, \boldsymbol{s}_j) = \left\|\boldsymbol{s}_i - \boldsymbol{s}_j\right\|$ is Euclidean distance. Let $\zeta^{\ell}(\boldsymbol{s}_i)$ denotes the neighborhood nodes set of the node $\boldsymbol{s}_i$, l is small, we have $\zeta^{\ell}(\boldsymbol{s}_i) = \left\{\boldsymbol{s}_j \middle| 0 < \left\|\boldsymbol{s}_i - \boldsymbol{s}_j\right\| \le \ell, \boldsymbol{s}_j \in S\right\}$. Then define non-neighborhood nodes set of the node $\boldsymbol{s}_i$, that is, $\overline{\zeta^{\ell}(\boldsymbol{s}_i)} = \left\{\boldsymbol{s}_j \middle| \ell < \left\|\boldsymbol{s}_i - \boldsymbol{s}_j\right\|, \boldsymbol{s}_j \in S\right\}$. The code of the rth dimension parameter in node $\boldsymbol{s}_i$ is $s_r^i = s_1^i s_2^i \cdots s_j^i \cdots s_{l_i}^i$, with j increasing, the weight coefficient decreases in the bit string, the j bit influences the parameter values with decreasing. Therefore, if two strings in the front bits are different, the two nodes to each other are their respective non-neighborhood nodes; only on the back order bits of the nodes are different, the two nodes to each other are their respective neighborhood nodes. The local search operator Ψ guides the exploration from $\boldsymbol{s}_i$ to the closest node $\boldsymbol{s}_i(k+1)$

which is in $\zeta^{\ell}(\boldsymbol{s}_i)$, and is denoted as $\boldsymbol{s}_i(k+1) \leftarrow \Psi(\boldsymbol{s}_i(k))$; according to a predefined probability, the random long-range search operator Γ selects a point $\boldsymbol{s}_i'(k)$ which is in $\overline{\zeta^{\ell}(\boldsymbol{s}_i)}$ as a candidate point to replace $\boldsymbol{s}_i(k)$, and is denoted as $\boldsymbol{s}_i'(k) \leftarrow \Gamma(\boldsymbol{s}_i(k))$. We divide the coding string $s_1^i s_2^i \cdots s_{l_i}^i$ into two sub-strings: $s_1^i s_2^i \cdots s_{l_i'}^i$ and $s_{l_i'+1}^i s_{l_i'+2}^i \cdots s_{l_i}^i$. The change of u bits in the sub-string $s_1^i s_2^i \cdots s_{l_i'}^i$ is random long-range search, and is denoted as Γ^u, $u \le l_i'$; the change of v bits in the sub-string $s_{l_i'+1}^i s_{l_i'+2}^i \cdots s_{l_i}^i$ is local search, and is denoted as Ψ^v, $v \le l_i - l_i'$. As the algorithm which based on the principle of small world phenomenon stresses on local search, therefore, $l_i' << l_i$. The implementation frequency of Ψ^v and Γ^u is influenced by the probability p_s, usually, p_s>0.5, which ensures more implementation of local search rather than random long-range search.

3.4 The Pseudo-code of DSWA

Based on the decimal-coding strategy, local search operator and random long-range search operator, the pseudo-code of DSWA is shown as the following:

```
program Decimal-coding Small World Algorithm(DSWA).
  const  n = 20; MaxIterations=50; l_i=12; P_s=0.8; N_i=4;
  var    k: 0..MaxIterations; j: 1..N_i;
  begin
    k := 0;
    while k< MaxIterations
       S'(k) ← S(k);
       for each s'(k) ∈ S'(k)
         repeat
           if P_s ≥ rand
              s'(k+1) ← Ψ(s'(k));
           else
              s'(k+1) ← Γ(s'(k));
           end if
           if  f(e^-1(s'(k+1))) < f(e^-1(s(k+1)))
              s'(k+1)  si(k+1);
           end if
           j := j+1
         until j = N_i
       end for
     k := k+1
    until k = MaxIterations
 end.
```

Where, rand∈ [0,1], $e^{-1}(\cdot)$ is the decoding function.

4 Experimental Studies

Picking equation (2) as the fitness function, a_1=0.999, a_2=0.001, a_3=2.0, a_4=100. The value of these coefficients indicate that the fitness function hope to restrain the overshoot than other optimization guidelines. As a comparison, a real encoding RGA [2], IPSOM [3] and DSWA are used for PID parameters tuning. In RGA, the size of individuals is 30, the crossover and mutation probability are P_c=0.9, P_m=0.033; In IPSOM, the number of particles is 30, accelerator coefficients c_1=c_2=2, inertia weight w_{ini}=0.9, w_{end}=0.4; In DSWA, the nodes number is 30, the encoding length of node is l_i=12, provisional transmission nodes number is 4. The termination condition of three algorithms is the maximum allowable iterations $k_{\max}$=50. The search space $K_p \in [0,2]$, $K_i \in [0,2]$, $K_d \in [0,2]$. Two typical high-order systems are seen as the objective systems, which transfer functions are given as equation (4) and (5).

$$G_1(s) = \frac{523500}{s^3 + 87.35s^2 + 10470s} \tag{4}$$

$$G_2(s) = \frac{0.1s + 10}{0.0004s^4 + 0.0454s^3 + 0.555s^2 + 1.51s + 11} e^{-0.2s} \tag{5}$$

Step signal is the input, the sampling time: T_{s1}=0.001s and T_{s2}=0.005s, 10 times simulations were carried out, and the best results were selected respectively. The step response characteristics of each algorithm are shown in Table 1 and Table 2. As can be seen from Table 1, the characteristics of system 1, the DSWA has the best performance than the other two algorithms. However, in Table 2, RGA-PID has better performance than the other algorithms, DSWA-PID has the smallest evaluation value. As a whole, DSWA is a feasible method to tune PID parameters.

Table 1. The characteristics of system 1 ($\Delta = 0.02$ **,** sampling time: T_{s1}=0.001s)

Algorithm	K_p	K_i	K_d	M_p	E_{ss}	t_r/s	t_s/s	Best values
IPSOM-PID	0.748	0.828	0.005	3.25%	0	0.085	0.125	53.371
RGA-PID	0.922	0.171	0.039	19.8%	0	0.014	0.165	79.815
DSWA-PID	0.864	0.375	0.021	0.04%	0	0.020	0.105	42.718

Table 2. The characteristics of system 2 ($\Delta = 0.02$ **,** sampling time: T_{s2}=0.005s)

Algorithm	K_p	K_i	K_d	M_p	E_{ss}	t_r/s	t_s/s	Best values
IPSOM-PID	0.057	1.852	0.148	10.1%	0	1.135	2.384	169.64
RGA-PID	0.088	1.649	0.120	3.25%	0	1.320	2.355	154.12
DSWA-PID	0.059	1.534	0.102	5.65%	0	1.575	3.252	153.43

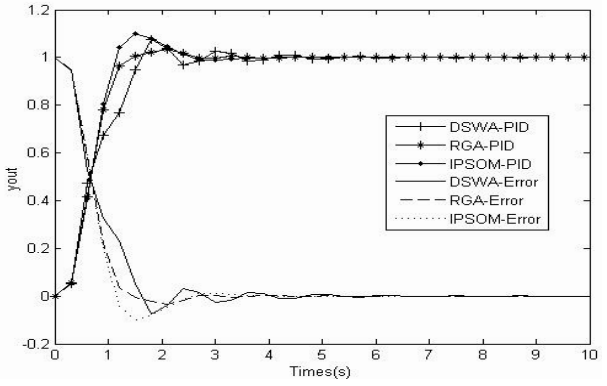


Fig. 2. Step response and error curves of system 2 with different controller

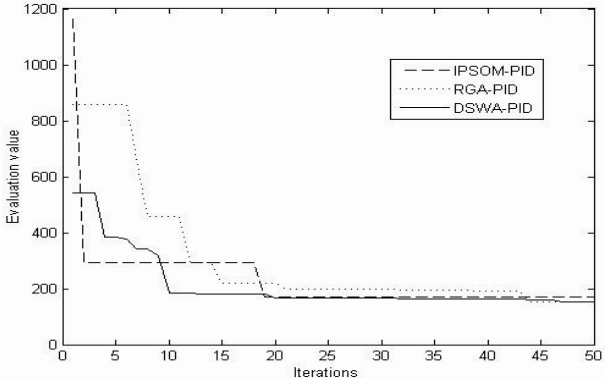


Fig. 3. The changes evaluation value by three algorithms (system 2)

Step response and their error curves as shown in Figure 2. Figure 3 shows the changes of the evaluation value by three algorithms. As we can see from Fig. 2 and Fig. 3, the DSWA-PID has better performance for the high-order system with time-delay, it can continuously search the global optima rather than trap in local optima. So the results indicate that using DSWA for tuning PID parameters is effective as RGA and IPSOM.

5 Concluding Remarks

The algorithm inspired by the small world phenomenon is used to tune PID parameters in this paper. Due to the PID control parameters optimization problem easily trapping in local optimum, by choosing reasonable evaluation criteria for control process, the PID parameters tuning with small world optimization process is presented. Then using decimal-coding mechanism to put forward a local search operator

and random long-range search operator, a decimal-coding small world optimization algorithm is proposed for PID parameters tuning, two typical systems are used to test for PID parameters tuning with IPSOM, RGA and DSWA. The results show that DSWA can achieve better evaluation value as RGA and IPSOM, which suggests that DSWA can overcome to convergence local optimum in a certain extent, and it has potential to solve the complex problems, such as PID parameters tuning.

Acknowledgements. This work is supported in part by the National Natural Science Foundation of China (70671083, 50505034).

References

1. Cominos, P., Munro, N.: PID Controllers: Recent Tuning Methods and Design to Specification. IEEE Proceedings Control Theory and Applications 149, 46–53 (2002)
2. Meng, X.Z., Song, B.Y.: Fast Genetic Algorithm Used for PID Parameter Optimization. In: IEEE Proceedings of International Conference on Automation and Logistics, pp. 2144–2148. IEEE Press, New York (2007)
3. Ren, Z.W., San, Y., Chen, J.F.: Improved Particle Swarm Optimization and its Application Research in Tuning of PID Parameters. Journal of System Simulation 18, 2870–2873 (2006)
4. Watts, D.: Six Degrees: the Science of a Connected Age. W.W. Norton & Company, New York (2004)
5. Du, H.F., Zhuang, J., Zhang, J.H.: Small-world Phenomenon for Function Optimization. Journal of Xi'an Jiaotong University 39, 1011–1015 (2005)
6. Liu, J.K.: Advanced PID Control and MATLAB Simulation. Publishing House of Electronics Industry, Beijing (2003)
7. Kleinberg, J.M.: The Small-World Phenomenon and Decentralized Search. SIAM News 37, 1–2 (2004)

SLAM by Combining Multidimensional Scaling and Particle Filtering

Hongmo Je[1] and Daijin Kim[2]

[1] Intelligent Multimedia Laboratory, Dept. of Computer Science & Engineering, Pohang University of Science and Technology (POSTECH), Pohang, Korea
invu71@postech.ac.kr

[2] Intelligent Multimedia Laboratory, Dept. of Computer Science & Engineering, Pohang University of Science and Technology (POSTECH), Pohang, Korea
dkim@postech.ac.kr

Abstract. This paper presents an algorithm for the simultaneous localization and mapping (SLAM) problem. Inspired by the basic idea of the FastSLAM which separates the robot pose estimation problem and mapping problem, we use the particle filter (PF) to estimate the pose of individual robot and use the multidimensional scaling (MDS), one of the distance mapping method, to find the relative coordinates of landmarks toward the robot. We apply the proposed algorithm to not only the single robot SLAM, but also the multi-robot SLAM. Experimental results demonstrate the effectiveness of the proposed algorithm over the FastSLAM.

1 Introduction

The first implemented SLAM using an Extended Kalman Filter (EKF) is introduced in [1]. In the EKF framework, a high-dimensional Gaussian is used to represent the robot pose and landmark position estimates. The key limitation of the EKF based SLAM is to require that features in the environment must be uniquely identifiable with the Gaussian noise assumption. If this requirement can not be satisfied, the EKF based SLAM will be failed [2]. Another popular family of SLAM algorithms is called FastSLAM [3]. FastSLAM algorithms were developed based on this property of the SLAM problems. Particle Filters (PFs) are used to estimate robot path.

MDS is a well-studied distance mapping method. Several MDS-based localization in sensor networks has already been presented in [4]. Multi-robot localization problem using the MDS also have been proposed in [5]. The strong merit of the MDS is that it needs only dissimilarity matrix to find relative position of objects. However, the MDS has not yet been applied to the SLAM problem, because it can not deal with the uncertainty of the observation, while the SLAM problem is needed to handle the noisy information, such as the motion controls and the observations. In this paper, we propose PF-MDS algorithm for SLAM. We use the particle filter to estimate the robot path with the motion model, and use the MDS to estimate the relative positions of the landmarks with the

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 825–832, 2008.
© Springer-Verlag Berlin Heidelberg 2008

measurement model. By combining the particle filter and the MDS, the proposed SLAM algorithm can handle the uncertainty of both robot motions and observations.

This paper is organized as follows. Section 2 describes related works on the MDS and FastSLAM. Section 3 introduces the PF-MDS SLAM in detail. The extension to the multi-robot SLAM including map fusing is explained in section 4. Section 5 presents the experimental results of the both single-robot and multi-robot SLAM simulation. Finally, section 6 presents the conclusion and the future work.

2 Related Works

2.1 Multidimensional Scaling

The task of classical MDS is to find the coordinates of p dimensional points $\mathbf{x}_r(r = 1, \ldots, n)$ from a Euclidean distance matrix. The Euclidean distance between $\mathbf{x}_r$ and $\mathbf{x}_s$ is defined as $d_{ij}^2 = (\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)$. Let the inner product of $\mathbf{x}$ is $\mathbf{B}$, where $[\mathbf{B}]_{ij} = b_{ij} = \mathbf{x}_i^T\mathbf{x}_j$. Then the Euclidean distance d_{ij}^2 can be represented by $\mathbf{B}$ as

$$d_{ij}^2 = \mathbf{x}_i^T\mathbf{x}_i + \mathbf{x}_j^T\mathbf{x}_j - 2\mathbf{x}_i^T\mathbf{x}_j = b_{ii} + bjj - 2b_{ij}. \tag{1}$$

By centering the coordinator matrix $\mathbf{X}$ to origin that($\sum_{i=1}^{b} b_{ij} = 0$) and summing Eq. (1) over i, over j, and over i and j, we find that

$$\frac{1}{N}\sum_{i=1}^{N} d_{ij}^2 = \frac{1}{N}\sum_{i=1}^{N} b_{ii} + b_{jj}, \frac{1}{N}\sum_{j=1}^{N} d_{ij}^2 = b_{ii} + \frac{1}{N}\sum_{j=1}^{N} b_{jj},$$

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij}^2 = \frac{2}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} b_{ii}. \tag{2}$$

By combining Eq. (1) and (2), we find that

$$b_{ij} = -\frac{1}{2}(d_{ij}^2 - d_{i\cdot}^2 - d_{\cdot j}^2 + d_{\cdot\cdot}^2) = (a_{ij} - a_{i\cdot} - a_{\cdot j} + a_{\cdot\cdot}),$$

where $a_{ij} = -\frac{1}{2}d_{ij}^2$. We can rewrite $\mathbf{B} = \mathbf{HAH}$ by double centering for the matrix of $[a_{ij}]$=$\mathbf{A}$, where the centering matrix $\mathbf{H} = \mathbf{I} - N^{-1}\mathbf{1}\mathbf{1}^T$. Since the matrix $\mathbf{B}$ is symmetric, positive semi definete, and $rank(\mathbf{B}) = rank(\mathbf{X}^T\mathbf{X}) = rank(\mathbf{X}) = p$, it can be decomposed as $\mathbf{B} = \mathbf{\Gamma\Lambda\Gamma}^T$, where $\mathbf{\Lambda} = diag(\lambda_1, \cdots, \lambda_p)$ is the diagonal matrix of $\mathbf{B}$, which corresponds to the eigen vector $\mathbf{\Gamma} = (\gamma_1, \cdots, \gamma_p)$. Hence, it can be represent the coordinators of $\mathbf{X} \in \mathbf{R}^p$ space as $\mathbf{X} = \mathbf{\Gamma\Lambda}^{1/2}$.

2.2 FastSLAM

Let us begin with explaining the basic of the SLAM problem. The map contains N landmarks is denoted $\Theta = [\theta_1\ \theta_2\ \ldots\ \theta_N]^T$. The path of robot is denoted

$\mathbf{x}^t = \mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_t$, where t is the time index. So $\mathbf{x}^t$ and $\mathbf{x}_t$ are different in that $\mathbf{x}_t$ means the pose of robot at time t. Most probabilistic approaches to SLAM have tried to solve the following posterior distribution: $p(\Theta, \mathbf{x}^t | \mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t)$, where $\mathbf{z}^t = \mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_t$ is a sequence of measurements, $\mathbf{u}^t = \mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_t$ is a sequence of robot controls, and $\mathbf{n}^t = \mathbf{n}_1, \mathbf{n}_2, ..., \mathbf{n}_t$ is a data association variable. In general, the data association variable $\mathbf{n}^t$ is considered as a known variable to make the problem easy.

In the FastSLAM, the posterior distribution can be factored as

$$p(\Theta, \mathbf{x}^t | \mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t) = p(\mathbf{x}^t | \mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t) \prod_{i=1}^{N} p(\theta_i | \mathbf{x}^t, \mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t). \tag{3}$$

This factorization states that the calculation of robot paths and maps can be decomposed into $N + 1$ probabilities. The real implementation of FastSLAM is based on the the rao-blackwellized particle filtering (RBPF) algorithm introduced in [6]. Each particle has the path of robot and its own map, consisting of N individual landmark estimates as follows.

$$\begin{aligned} \varphi_t^{[1]} &= (\mathbf{x}^{t,[1]}, \mu_{1,t}^{[1]}, \Sigma_{1,t}^{[1]}, \cdots, \mu_{N,t}^{[1]}, \Sigma_{N,t}^{[1]}) \\ &\vdots \\ \varphi_t^{[M]} &= (\mathbf{x}^{t,[M]}, \mu_{1,t}^{[M]}, \Sigma_{1,t}^{[M]}, \cdots, \mu_{N,t}^{[M]}, \Sigma_{N,t}^{[M]}), \end{aligned} \tag{4}$$

where $[k]$ is the index of the particle, $\mathbf{x}^{t,[k]}$ is the robot path in the k-th particle, and $\mu_{i,t}^{[2]}, \Sigma_{i,t}^{[k]}$ are the mean and covariance representing the landmark θ_i attached at the k-th particle.

3 The Proposed Method

In this section, we explain the proposed method in detail. We combines particle filtering and MDS to solve the SLAM problem. The particle filter is used to estimate the robot path, and the MDS is used to estimate the position of the landmarks. More details on the proposed method are described in [7].

3.1 Robot Path Estimation

We use the particle filter to estimate the posterior of the robot path $p(\mathbf{x}^t | \mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t)$. The posterior is denoted by the set of particle as χ_t and each particle $\mathbf{x}_t^{[k]} \in \chi_t$ is represent as $\chi_t = \{\mathbf{x}^{t,[k]}\} = \{\mathbf{x}_1^{[k]}, \mathbf{x}_2^{[k]}, ..., \mathbf{x}_t^{[k]}\}_{\text{for } k=1,...,M}$, where the superscript k refers to the k-th particle in the set and M is the number of particles. Then, the new pose of the robot at time t is sampled from the motion model $\mathbf{x}_t^{[k]} \sim p(\mathbf{x}_t | \mathbf{u}_t, \mathbf{x}_{t-1}^{[i]})$, where $\mathbf{u}_t$ is the most recent motion command. Then, a temporary set of particles is generated by adding this sampled pose into the previous path $\mathbf{x}^{t-1,[k]}$. Assuming the prior set of particles in χ_{t-1} is distributed according to $p(\mathbf{x}^{t-1} | \mathbf{z}^{t-1}, \mathbf{u}^{t-1}, \mathbf{n}^{t-1})$, where $\mathbf{n}^{t-1}$ is the data association variable, the new

temporary set of particles is distributed according to $p(\mathbf{x}^t|\mathbf{z}^{t-1}, \mathbf{u}^t, \mathbf{n}^{t-1})$. This distribution is the *proposal distribution* of our proposed particle filter based robot path estimation.

After generating all temporary particles, the new set of particles χ_t is generated by sampling from this temporary particle set. Each particles $\mathbf{x}^{t,[k]}$ is resampled with a probability proportional to a weight or importance $w_t^{[k]}$ which is defined as $w_t^{[k]} = \frac{f(\mathbf{x})}{g(\mathbf{x})}$, where $f(\mathbf{x})$ is the target distribution and $g(\mathbf{x})$ is the proposal distribution. By resampling, the new set of particles χ_t must be distributed according to an approximation to the desired posterior $p(\mathbf{x}^t|\mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t)$. By using Bayesian theorem and the Markov property , the importance weights $w_t^{[k]}$ for resampling from the robot path particles can be approximated as follows.

$$\begin{aligned} w_t^{[k]} &\propto \frac{p(\mathbf{x}^{t,[k]}|\mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t)}{p(\mathbf{x}^{t,[k]}|\mathbf{z}^{t-1}, \mathbf{u}^t, \mathbf{n}^{t-1})} \\ &= \frac{p(\mathbf{z}_t, \mathbf{n}_t|\mathbf{x}^{t,[k]}, \mathbf{z}^{t-1}, \mathbf{u}^t, \mathbf{n}^{t-1})}{p(\mathbf{z}_t, \mathbf{n}_t|\mathbf{z}^{t-1}, \mathbf{u}^t, \mathbf{n}^{t-1})} : \text{Bayes theorm} \\ &\propto p(\mathbf{z}^t, \mathbf{n}^t|\mathbf{x}^{t,[k]}, \mathbf{z}^{t-1}, \mathbf{u}^t, \mathbf{n}^{t-1}) \\ &\approx \int p(\mathbf{z}_t|\theta_{\mathbf{n}_t}^{[k]}, \mathbf{x}_t^{[k]}, \mathbf{n}_t) p(\theta_{\mathbf{n}_t}^{[k]}) d\theta_{\mathbf{n}_t}, \end{aligned} \tag{5}$$

where $\theta_{\mathbf{n}_t}$ is the position of landmarks which is identified by using the data association with the current observations. $p(\mathbf{z}_t|\theta_{\mathbf{n}_t}^{[k]}, \mathbf{x}_t^{[k]}, \mathbf{n}_t)$ can be approximated by the landmark position estimation described in 3.4. Here, $p(\theta_{\mathbf{n}_t}^{[k]})$ is the Gaussian posterior.

3.2 MDS Based Relative Localization

Let us consider the relative localization between a robot and observed landmarks. The position vector comprising robot pose, $\mathbf{x}_r$, and the observed landmarks, $\mathbf{x}_m$, is given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_r \\ \Theta \end{bmatrix} = [x \ y \ \alpha \ \theta_{1x} \ \theta_{1_y} \ ... \ \theta_{Lx} \ \theta_{Ly}]^T, \tag{6}$$

where x, y, and α denote the robot pose, and θ_{ix}, θ_{iy} are the coordinates of i-th landmark, for $i = 1, ..., L$, where L denote the number of the landmarks in the current observation. Because the Gram matrix is the positive semidefinite matrix, it is possible to calculate $\mathbf{V} = \mathbf{\Gamma}\mathbf{\Lambda}^{1/2}$, after evaluating the singular value decomposition of $\mathbf{K} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}^T$.

3.3 Data Association

In a real world SLAM, it is difficult to recognize the natural landmarks while the artificial landmarks can be identified easily. The data association is a problem to find the relationship between the current observation $\mathbf{z}_t$ and the set of landmark positions Θ. It can also determine whether the observation of landmark is new,

it means previously unseen, or explored landmark. If the observed landmark is new, it should be added into the set of landmark immediately.

Similar to the FastSLAM, our proposed PF-MDS algorithm performs the data association for each particle. Thus, instead of finding the unique data association variable $\mathbf{n}_t$ in the EKF-SLAM, the PF-MDS estimates the following equation.

$$\mathbf{n}_t^{[k]} = \arg\max_{\mathbf{n}_t} p(\mathbf{z}_t|\mathbf{x}_t^{[k]}, \mathbf{n}_t). \tag{7}$$

3.4 Landmark Positions Estimation

After estimating the robot path, the estimation of landmark positions is needed. We use the the measurement model, $\mathbf{z}_t$, having the range and bearing from a robot toward landmark, which is defined as follows.

$$\begin{aligned} R_i &= \sqrt{(x_r - \theta_{ix})^2 + (y_r - \theta_{iy})^2} + \varepsilon_r \\ , \phi_i &= tan^{-1}\left(\frac{x_r - \theta_{ix}}{y_r - \theta_{yx}}\right) - \alpha + \varepsilon_\phi, \end{aligned} \tag{8}$$

where (x_r, y_r) is the pose of the robot and α is its bearing,$(\theta_{ix}, \theta_{iy})$ is the position of the landmark, and $\varepsilon_r \sim N(0, \sigma_r)$ and $\varepsilon_\phi \sim N(0, \sigma_\phi)$ are the noise associated with the range and angle measurement. In order to obtain particles of landmark positions, the relative locations of the current observed landmarks toward the robot is first generated by using MDS. By adding the uncertainty of the range measurement, σ_r, and the uncertainty of the anlge measurement, σ_ϕ, N particles of the current observed landmark positions are generated. We must consider the posterior over the i-th landmarks pose in case whether the landmark is observed or not. If the landmark is observed , it means $i = \mathbf{n}_t$ and hence $\theta_i = \theta_{\mathbf{n}_t}$, the posterior is defined as

$$p(\theta_{i=\mathbf{n}_t}|\mathbf{x}^t, \mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t) \propto p(\mathbf{z}_t|\theta_{\mathbf{n}_t}, \mathbf{x}_t, \mathbf{n}_t) p(\theta_{\mathbf{n}_t}|\mathbf{x}^{t-1}, \mathbf{z}^{t-1}, \mathbf{u}^{t-1}, \mathbf{x}^{n-1}). \tag{9}$$

The importance weights for resampling are proportional to the following ratio.

$$w_t^{[l]} \propto \frac{p(\theta_{\mathbf{n}_t}^{[l]}|\mathbf{x}^t, \mathbf{z}^t, \mathbf{u}^t, \mathbf{n}^t)}{p(\theta_{\mathbf{n}_t}^{[l]}|\mathbf{x}^{t-1}, \mathbf{z}^{t-1}, \mathbf{u}^{t-1}, \mathbf{n}^{t-1})} \propto p(\mathbf{z}_t|\theta_{\mathbf{n}_t}^{[l]}, \mathbf{x}_t, \mathbf{n}_t).$$

$p(\mathbf{z}_t|\theta_{\mathbf{n}_t}^{[l]}, \mathbf{x}_t, \mathbf{n}_t)$ can be easily calculated using a likelihood function. The following equation introduced in [8] is used as the likelihood function for the given observations.

$$L(\mathbf{z}_t|\theta_{\mathbf{n}_t}, \mathbf{x}_t, \mathbf{n}_t) \triangleq \frac{1}{\sqrt{2\pi(\sigma_r^2 + \sigma_\phi^2)}} \exp\left(\frac{-(R - \hat{R})^2}{\sigma_r^2} + \frac{-(\phi - \hat{\phi})^2}{\sigma_\phi^2}\right), \tag{10}$$

where $\hat{R}$, $\hat{\phi}$ are the expected range and angle for a particle.

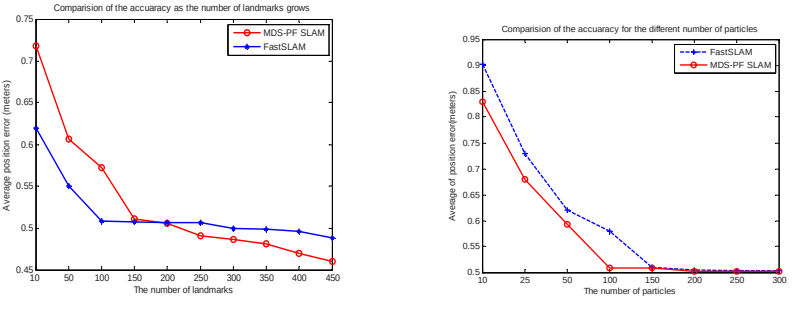
4 Multi-robot SLAM

The multi-robot SLAM is a problem that several robot seek to build a joint map when they explore same environment. The cooperation of multiple robots helps them solve the SLAM problem more efficiently [9]. For map fusing, we use the Maximum A Posterior (MAP) estimation of the position of landmarks from individual global map.

5 Experimental Results

5.1 Case I: Single Robot SLAM

It is meaningful to compare the performance between the FastSLAM and the proposed SLAM. The space of the robot exploring is 150 meters $\times$ 150 meters. The number of landmarks are from 10 to 450. The number of particles are from 10 to 300. In general, the number of landmarks and the number of particles have an effect on the performance of SLAM. Fig. 1(a) indicates the result when the number of particles are fixed as 100 and the number of landmarks grows. Fig. 1(b) indicates the result when the number of landmarks are fixed as 200 and the number of particles are alternated. From these results, we recognize the proposed SLAM is almost match with the FastSLAM in terms of the performance of position estimation. Moreover, the proposed SLAM needs less particles than than the FastSLAM.



(a) (b)

Fig. 1. (a) Comparison of the number of landmarks (when 100 particles)., (b) Comparison of the number of particles (when 200 landmarks).

5.2 Case II: Multi-robot SLAM

We also demonstrate the performance of the proposed PF-MDS based multi-robot SLAM. Three types of the simulation environment has been defined to verify the efficiency of multi-robot SLAM as the size of exploring space grows. We defined four scenarios to demonstrate the effect of multi-robot SLAM as follows.

– SCN-1 : One Loop circulation with Single Robot (OLSR)
– SCN-2 : One Loop circulation with Multiple Robots (OLMR)
– SCN-3 : Multiple Loop circulations with Single Robot(MLSR)
– SCN-4 : Multiple Loop circulations with Multiple Robots (MLMR)

Each robot in simulator has 0.3 meter-sized wheel base with maximum speed of 1.2m/s and is equipped with a range-bering sensor (laser scanner) with the maximum range of 20 meter and a 180 degrees of frontal-angle-view. The control noise is 0.1 meters in driving distance and 3 degrees in turning angle. The uncertainty of measurement is 0.3 meters in range and 5 degrees in angle. The control signal is generated at every 25 mili-seconds and the measurement observations are obtained at every 200 mili-seconds. From 1 to 10 robots are put into the simulation environment. The experiment is executed 100 times with different starting points for each scenario and the simulation environment. We initially set the number of particles to 50, however, the average number of particles becomes less than 10 by applying the selective resampling method, which finds the effective number of particles, introduced in [10].

Table 1 summarizes the best results of the experiment, where the 'NUMS' is the number of robots when the accuracy is highest, 'LOOPS' is the iterations of loop circulation when the accuracy is highest, and 'MSE' is the mean of square errors between the ground truth position and the estimated position.

Table 1. The best results of the experiment

Environment	Scenario	MSE(meters)	NUMS	LOOPS
Small-Sized	SCN-1	0.79	*	*
	SCN-2	0.71	3	*
	SCN-3	0.58	*	4
	SCN-4	0.51	2	3
Medium-Sized	SCN-1	0.89	*	*
	SCN-2	0.79	6	*
	SCN-3	0.68	*	6
	SCN-4	0.60	5	5
Large-Sized	SCN-1	1.54	*	*
	SCN-2	1.39	10	*
	SCN-3	1.08	*	4
	SCN-4	0.95	10	5

6 Conclusions and the Future Works

This paper proposes a SLAM algorithm by combining particle filtering and MDS. We think the SLAM problem as the factorized estimation of both the robot path and the position of landmarks. Therefore, the proposed PF-SLAM can be regarded as a member of the FastSLAM family. Like the FastSLAM, we use the particle filter to estimate the robot path with motion model. Unlike the FastSLAM, we use the MDS based relative localization, instead of the EKF, to

generate the particle of landmarks. We also show that the proposed PF-MDS can be easily extended to multi-robot SLAM problem. Although we use a simple strategy for map fusion, the experimental results demonstrate the effectiveness of the multi-robot SLAM based on PF-MDS. To obtain more impressive advantage of multi-robot SLAM, it is essential to explore into the advanced map fusion and cooperative data association methodology.

Ackowledgements. The authors would like to thank the Ministry of Education of Korea for its financial support toward the Division of Computer Science&Engineering at POSTECH through BK21 program, and was supported by the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy (MOCIE).

References

1. Moutarlier, P., Chatila, R.: Stochastic multisensory data fusion for mobile robot location and environment modeling. In: 5th Int. Symposium on Robotics Research, Tokyo (1989)
2. Hu, W., Downs, T., Wyeth, G., Milford, M., Prasser, D.: A modified particle filter for simultaneous robot localization and landmark tracking in an indoor environment. Journal of Intelligent and Robotic Systems 46, 365–382 (2006)
3. Montemerlo, M., Thrun, S.: Simultaneous localization and mapping with unknown data association using fastslam. In: ICRA 2003, PA, USA, vol. 2, pp. 1985–1991 (2003)
4. Klock, H., Buhmann, J.M.: Multidimensional scaling by deterministic annealing. In: Energy Minimization Methods in Computer Vision and Pattern Recognition, pp. 245–260 (1997)
5. Je, H., Sukhatme, G.S., Kim, D.: Partially observed distance mapping for cooperative multi-robot localization. Journal of Intelligent Service Robot (JISR) (to be appeared, 2008)
6. Murphy, K.: Bayesian map learning in dynamic environments. In: NIPS 1999 (1999)
7. Je, H., Kim, D.: Pf-slam: An algorihm of combining multidimensional scaling and particle filter for simultaneous multi-robot localization and mapping. Technical Report, TR-POSTECH-IM2008-PFMDS (2008)
8. Dissanayake, N.P., Clark, S., Durrant-Whyte, H.F., Csorba, M.: A solution to the simultaneous localization and map building (slam)problem. IEEE Transactions on Robotics and Automation 17, 229–241 (2001)
9. Howard, A.: Multi-robot simultaneous localization and mapping using particle filters. International Journal of Robotics Research 25, 1243–1256 (2006)
10. Grisetti, G., Stachniss, C., Burgard, W.: Improving grid-based slam with raoblackwellized particle filters by adaptive proposals and selective resampling (2005)

Application of Ant Colony System to an Experimental Propeller Setup

Jih-Gau Juang and Yi-Cheng Lu

Department of Communication and Guidance Engineering
National Taiwan Ocean University, Keelung, Taiwan
jgjuang@mail.ntou.edu.tw1

Abstract. This paper presents a control problem involving an experimental propeller setup that is called the twin rotor multi-input multi-output system (TRMS). The control objective is to make the beam of the TRMS move quickly and accurately to the desired attitudes, both the pitch angle and the azimuth angle in the condition of decoupling between two axes. It is difficult to design a suitable controller because of the influence between two axes and nonlinear movement. For easy demonstration in the vertical and horizontal planes separately, the TRMS is decoupled by the main rotor and the tail rotor. An intelligent control scheme which utilizes a hybrid ACS-PID controller is implemented in the system. Simulation results show that the new approach can improve the tracking performance and reduce control force in the TRMS.

Keywords: MIMO system, ant colony system, PID control.

1 Introduction

PID controller is a well-known method for industrial control processes because of its simple structure. It contains three parameters: proportional term (Kp), integral term (Ki), and derivative term (Kd). To obtain optimal set of parameters is usually very difficult, sometimes good experience is needed. It is one of the most realistic subjects in industrial control processes to get the best set of parameters. There are many methods for tuning controller parameters such as Ziegler-Nichols (ZN) method [1], Internal Model Control (IMC) [2], and pole placement method [3]. Recently, intelligent systems such as genetic algorithm (GA) [4], particle swarm optimization [5], and ant system [6] have been applied to optimization problems. These systems can be utilized in searching optimal controller parameters.

Ant system (AS) was first proposed by Dorigo in 1991. It imitates the nature ant behavior of colony cooperation to look for the food. The ants will deposit pheromones on their route, then utilize path tracking ability to find the shortest path between their nest and a food source. In 1996, Dorigo proposed ant colony system (ACS) to solve the optimization problems and applied it to the classical Traveling Salesman Problem (TSP) [7]. Here we use parameter searching capability of the ACS with a conventional PID controller as a solution to the TRMS control problem. The parameters of the controller are tuned by the ACS that fitness function is formed by system performance index [8]. The system performance index is an optimization criterion for parameters

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 833–840, 2008.
© Springer-Verlag Berlin Heidelberg 2008

tuning of control system. It deals with a modification of the known integral of time multiplied by squared error criterion (ITSE) [9].

In previous work, PID controller is incapable to reduce oscillation and undershoot on horizontal position and overshoot and rising time on vertical position. These disadvantages lead to TRMS incompetent to track desired trajectory accurately. Control energy is also not optimized. In here, a PID controller is utilized, which combined with the ACS. The strict model of mathematics is needless in this paper with respect to the searching parameters of controller which is applied to the ACS. For this reason the design process is more concise and the controller is more robust to the traditional controller. The structure of software and hardware is to employ Matlab Simulink, VisSim, and I/O card, which can construct software simulation and PC-based control system. According to simulation results, the new approach has better performances in setpoint and path tracking control.

2 Twin Rotor MIMO System

The TRMS, as shown in Fig. 1, is characterized by complex, highly nonlinear and inaccessibility of some states and outputs for measurements, and hence can be considered as a challenging engineering problem [10]. The control objective is to make the beam of the TRMS move quickly and accurately to the desired attitudes, both the pitch angle and the azimuth angle in the condition of decoupling between two axes. The TRMS is a laboratory set-up for control experiment and is driven by two DC motors. Its two propellers are perpendicular to each other and joined by a beam pivoted on its base that can rotate freely in the horizontal and vertical plane. The joined beam can be moved by changing the input voltage to control the rotational speed of these two propellers. There is a pendulum counter-weight hanging on the joined beam which is used for balancing the angular momentum in steady state or with load. In certain aspects its behavior resembles that of a helicopter. It is difficult to design a suitable controller because of the influence between two axes and nonlinear movement. From the control point of view it exemplifies a high order nonlinear system with significant cross coupling. For easy demonstration in the vertical and horizontal separately, the TRMS is decoupled by the main rotor and tail rotor.

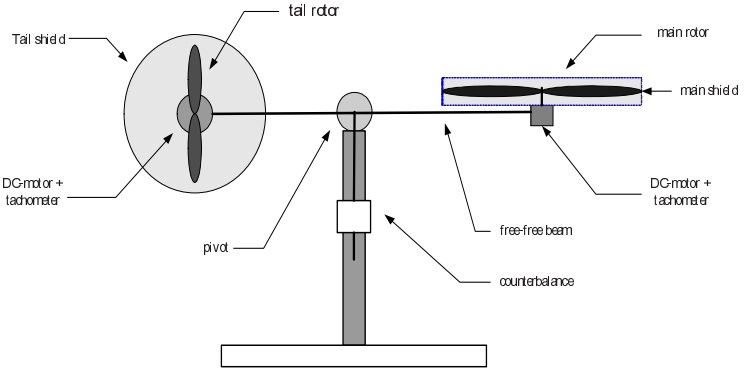


Fig. 1. Twin rotor multi-input multi-output system

3 Ant Colony System

ACS is an evolutionary heuristic algorithm that has been applied successfully to solve various combinatorial optimization problems. In this section, we will describe how to use the ACS algorithm to search for the global optimal values of the three PID adjustable parameters. In order to raise the efficiency of the ACS algorithm in solving the given optimization problem, some modifications to the ACS algorithm proposed in [11] are made in our research, including generating a clear motion environment for the ants and expressing the nine adjustable parameters using an ant moving path. When using the ACS algorithm to solve an optimization problem, the first step is to design a suitable motion environment for ants. We define that each parameter has four valid digits, as shown in Fig. 2. We express the value of KP, KI and KD on X-Y plane. First we draw twelve lines, L1, L2,…, L12, which have equal length and equal interval and are perpendicular to axis X. The ant proceeds from the origin O, moves to any node of next line, then reaches L12, this completes a tour. Each line only passes once. L1~L4, L5~L8, and L9~L12 represent the value of KP, KI , and KD, respectively. Each line has 10 points represent the values of 0 to 9. We use nij to denote the node j on line Li. The y coordinate of the node nij is denoted by yij. Obviously, the values of the parameters KP, KI, and KD are represented by the following formulas:

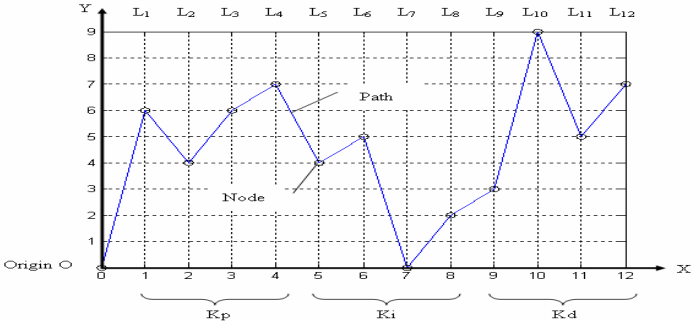


Fig. 2. Diagram of generating nodes and paths

$$\begin{cases} K_{\mathrm{P}} = y_{1,j} \times 10^{-1} + y_{2,j} \times 10^{-2} + y_{3,j} \times 10^{-3} + y_{4,j} \times 10^{-4} \\ K_{\mathrm{I}} = y_{5,j} \times 10^{-1} + y_{6,j} \times 10^{-2} + y_{7,j} \times 10^{-3} + y_{8,j} \times 10^{-4} \\ K_{\mathrm{D}} = y_{9,j} \times 10^{-1} + y_{10,j} \times 10^{-2} + y_{11,j} \times 10^{-3} + y_{12,j} \times 10^{-4} \end{cases} \tag{1}$$

The ant selects the next node n_{ij} among a candidate list to move to, which is based on the following transition rule:

$$j = \begin{cases} \arg\max_{u \in A} \{[\tau_{iu}(t)] \cdot [\eta_{iu}]^{\beta}]\} \ , & if \ \ q \le q_0 \\ J & , \ if \ \ q > q_0 \end{cases} \tag{2}$$

where A represents the set: {0, 1, …, 9}, τ_{iu}(t) is the pheromone concentration of node n_{iu}, η_{iu} represents the visibility of node niu, β is a parameter which controls the relative

importance of visibility η_{iu} versus pheromone concentration $\tau_{iu}(t)$, q is a random variable uniformly distributed over [0, 1], q0 is a tunable parameter (0<q0<1), and J is a node which is selected according to the following probability formula in (3)

$$p_{iJ}^{k}(t)=\frac{[\tau_{iJ}(t)]\cdot[\eta_{iu}]^{\beta}}{\sum_{w=0}^{9}[\tau_{iJ}(t)]\cdot[\eta_{iu}]^{\beta}} \tag{3}$$

Each ant modifies the environment in two different ways:

1). Local pheromone updating

As ant passes through a node, it updates the amount of pheromone by the following formula

$$\tau_{ij}(t)\leftarrow(1-\rho)\cdot\tau_{ij}(t)+\rho\cdot\tau_{0} \tag{4}$$

where ρ is the evaporation constant and $\tau_0=(ITAE_0)^{-1}$.

$$ITAE=\int_{0}^{t}t\cdot\left|e(t)\right|dt \tag{5}$$

where ITAE is the fitness value of the system when the gains KP, KI, and KD of the PID controller are tuned by the Ziegler-Nichols criterion.

2). Global pheromone updating

When all ants have completed a tour, the best tour needs to be updated using the global pheromone updating formula

$$\tau_{ij}(t)\leftarrow(1-\rho)\cdot\tau_{ij}(t)+\rho\cdot\Delta\tau_{ij}(t) \tag{6}$$

$$\Delta\tau_{ij}(t)=1/ITAE^{*} \tag{7}$$

where ITAE * is the value of ITAE corresponding to the best path T*.

4 Simulations

The TRMS is decoupled into vertical plane and horizontal plane. Each plane contains a PID controller. In order to control TRMS to reach a setpoint and track a specified trajectory quickly and accurately, the parameters of PID controller are tuned by the ACS. The target of the ACS is the parameters of PID controller KP, KI, and KD. The system performance index is used for fitness function that fit the output requirement. In the process of ACS, number of ants is 10, tunable parameter q0 is set to 0.85, pheromone decay coefficient ρ is set to 0.1 and adjustable β is set to 1.

When all ants finish traveling, update pheromone concentration. And then evaluate fitness value of each path. In the simulation, reference input in the horizontal plane is 1 rad and initial condition is 0 rad; reference input in the vertical plane is 0.2 rad and initial condition is -0.5 rad. Simulation time is from 0 to 50 seconds, computation procedure is in Fig. 3. The target parameters in the horizontal plane are gain factors of PID controller that are KP, KI and KD within range [0,1]. System performance requirements are that maximum overshoot is less than 1%, rise time, delay time, and steady state time are as short as possible. Fitness function in the ACS is

$$FitnessFunction=\frac{\kappa}{I_{PITSE}} \tag{8}$$

where κ is a constant, I_{PITSE} is the system performance index. Tuning results are in Tab.1. Simulation results of setpoint control are: total error is 23.4986, total control force is 7.9852, where simulation time is 50 seconds, sampling time is 0.5 seconds. Fitness value and step response are shown in Fig. 4 and Fig. 5. Simulations of tracking control with sine wave and square wave are shown in Fig. 6. The amplitudes of sine wave and square wave are 0.5 rad and 1 rad, respectively. Compare with the conventional PID control, as shown in Tab.2, the control force reduces more than 30% in sine wave response and 70% in square wave response. The new approach reduces tracking error and converges to the steady state faster than the PID control.

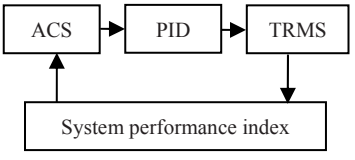


Fig. 3. ACS computation procedure

Table 1. Controller parameters in the horizontal plane

KPhh	KIhh	KDhh
0.4835	0.0000	0.3628

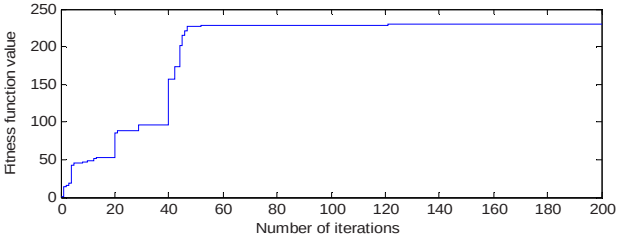


Fig. 4. Fitness value in the horizontal plane

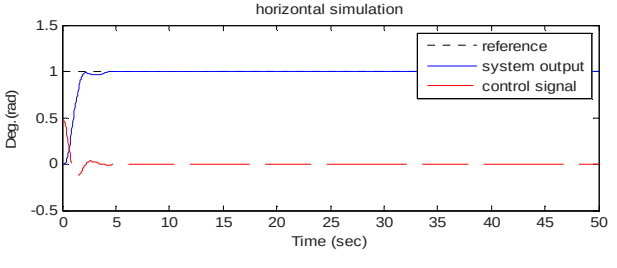


Fig. 5. Step response in the horizontal plane

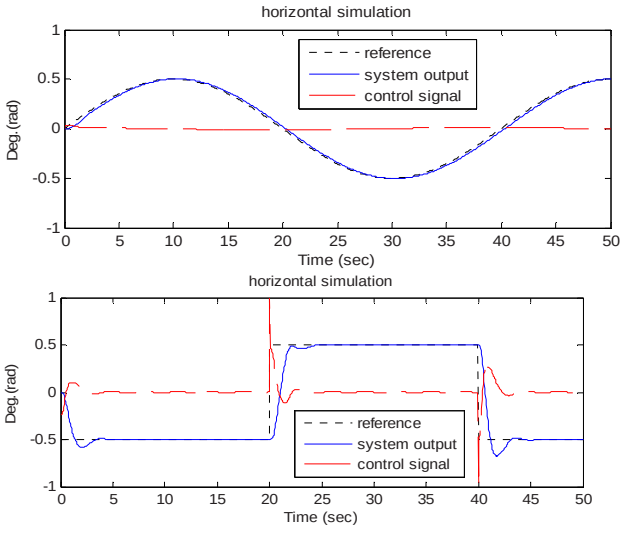


Fig. 6. Sine wave response and square wave response in the horizontal plane

Table 2. Comparison of total error and control force in the horizontal plane

	setpoint		sine wave		square wave	
	error	control	error	control	error	control
ACS	23.49	7.98	16.08	7.86	51.29	23.20
PID	26.21	17.68	11.58	12.47	86.36	78.66

Table 3. Controller parameters in the vertical plane

KPhh	***KIhh***	***KDhh***
0.0715	0.2861	0.1436

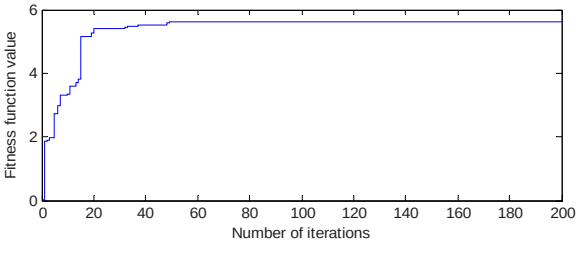


Fig. 7. Fitness value in the vertical plane

The target parameters in the vertical plane are the same as in the horizontal plane. Tuning results are in Tab. 3. Simulation results of setpoint control are: total error is 53.1482, total control force is 591.0788. Fitness value and step response are shown in

Fig. 7 and Fig. 8. Simulations of tracking control with sine wave and square wave are shown in Fig. 9. The amplitudes of sine wave and square wave are both 0.2 rad. Compare with the conventional PID control, as shown in Tab. 4, the new approach converges faster and has better performance.

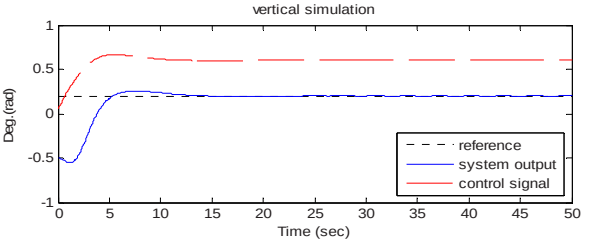


Fig. 8. Step response in the vertical plane

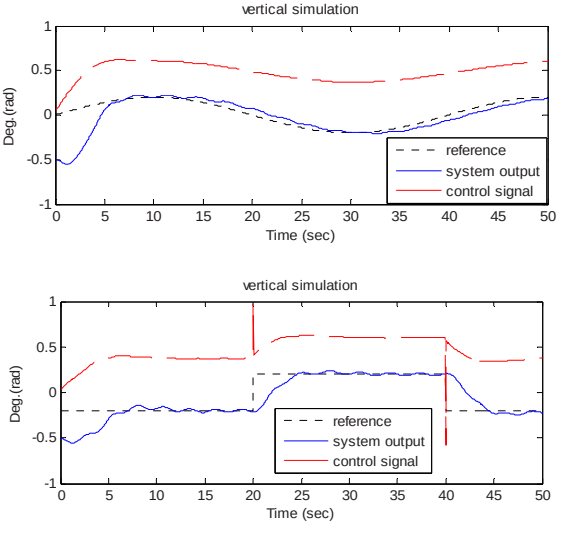


Fig. 9. Sine wave response and square wave response in the vertical plane

Table 4. Comparison of total error and control force in the vertical plane

	setpoint		sine wave		square wave	
	error	control	error	control	error	control
ACS	53.14	591.07	78.20	485.37	76.90	451.53
PID	67.84	687.53	72.69	1185.7	137.27	1234.8

5 Conclusion

In this paper, we investigate a control problem involving an experimental propeller setup that is called the twin rotor multi-input multi-output system (TRMS).

Experiments with setpoint and trajectory tracking control in the condition of decoupling between two axes are presented. PID controller is utilized to the TRMS. In order to design an optimal set of PID control gains, we apply ant colony system with system performance index to search the parameters of the controller that can make the TRMS track a desired trajectory or reach a setpoint quickly and accurately. In this paper, a specific system performance index is used for fitness function which focuses on time domain response of the TRMS. The simulation results show that the new approach can improve the positioning and tracking performances and reduce control force in the TRMS control problem.

Acknowledgement. This work was supported by the National Science Council, Taiwan, ROC under Grant NSC 95-2221-E-019 -054.

References

1. Ziegler, J.G., Nichols, N.B.: Optimum Setting for Automatic Controllers. Trans. ASME, 759–768 (1942)
2. Garcia, C.E., Morari, M.: Interanl Model Control. 1.A Unifying Review and Some New Results. Ind. Eng. Chem. Process. Dev. 21, 308–323 (1982)
3. Astrom, K.J., Hagglud, T.: PID Controller: Theory, Design and Tuning. Research Triangle Park, NC (1995)
4. Man, K., Tang, F.K.S., Kwong, S.: Genetic Algorithms. Springer, Heidelberg (1999)
5. Zheng, Y.L., Ma, L., Zhang, L., Qian, J.: On the Convergence Analysis and Parameter Selection in Particle Swarm Optimization. In: 2nd IEEE International Conference on Machine Learning and Cybernetics, pp. 1802–1807 (2003)
6. Dorigo, M., Maniezz, V., Colorni, A.: Ant System: An Autocatalytic Optimizing Process. Technical Report No. 91-016 Revised, Politecnico Di Milano, Italy (1991)
7. Dorigo, M., Maniezz, V., Colorni, A.: Ant System: Optimization by a Colony of Cooperating Agents. IEEE Transactions on System, Man and Cybernetics-Part B 26(1) (1996)
8. Boblan, D.-l.I.: A New Integral Criterion Parameter Optimization of Dynamic System with Evolution Strategy. VDI Berichte Nr.1526, Computational Intelligence, pp. 143–150 (2000)
9. Liu, W.K.: Applications of Hybrid Fuzzy-GA Controller and FPGA Chip to TRMS, M.S. thesis, Dept. of Comm. and Guidance Eng., National Taiwan Ocean Univ., Taiwan (2005)
10. Feedback Co, Twin Rotor MIMO System 33-220 User Manual, 3-000M5 (1998)
11. Zeng, Q., Tan, G.: Modified ACS Algorithm-Based Nonlinear PID Controller and Its Application to Cip-I Intelligent Leg. In: International Conference on Intelligent Computing (2007)

Bacterial Foraging Based Optimization Design of Fuzzy PID Controllers

Hung-Cheng Chen

Department of Electrical Engineering, National Chin-Yi University of Technology,
35, Lane 215, Sec. 1, Chungshan Road, Taiping, Taichung 411, Taiwan
hcchen@ncut.edu.tw

Abstract. In this paper, a bacterial foraging optimization scheme (BFOS) is proposed for the multi-objective optimization design of a fuzzy PID controller and applies it to the control of an active magnetic bearing (AMB) system. Different from PID controllers with fixed gains, the fuzzy PID controller is expressed in terms of fuzzy rules whose rule consequences employ analytical PID expressions. The PID gains are adaptive and the fuzzy PID controller has more flexibility and capability than the conventional ones. Moreover, it can be easily utilized to develop a precise and fast control algorithm in optimal design. The BFOS is used to design the fuzzy PID controller. The centers of the triangular membership functions and the PID gains for all fuzzy control rules are selected as parameters to be determined. The dynamic model of AMB system for axial motion is also presented. The simulation results of this AMB system show that a fuzzy PID controller designed via the proposed BFOS has good performance.

Keywords: Bacterial foraging optimization, Active Magnetic Bearing, Fuzzy PID Controller.

1 Introduction

In the past decades, conventional PID controllers are widely applied in industry process control. This is mainly because PID controllers have simple control structures, and are simple to maintain. [1,2]. To design such a controller, the proportional gains, the integral gains, and the derivative gains must be determined. However, a conventional PID controller may have poor control performance for nonlinear and/or complex systems that have no precise mathematical models. Various types of modified traditional PID controllers such as auto-tuning and adaptive PID controllers were developed to overcome these difficulties [3,4]. Since the PID gains are fixed, the main disadvantage is that they usually lack in flexibility and capability. Recently, many researchers attempted to combine conventional PID controllers with fuzzy logic [5,6]. Despite the significant improvement of these fuzzy PID controllers over their conventional counterparts, it should be noted that they still have some disadvantages. For example, the locations of the peak of the membership functions are fixed and not adjustable, and the fuzzy control rules are handed-designed rules.

To overcome the weaknesses mentioned above, we propose a multi-objective optimization method for the parameter tuning of fuzzy PID controllers based on a

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 841 – 849, 2008.
© Springer-Verlag Berlin Heidelberg 2008

bacterial foraging optimization scheme (BFOS) to solve the control problem of an AMB system. In this scheme, the foraging behavior of E. coli bacteria present in our intestines is mimicked. They undergo different stages such as chemotaxis, swarming, reproduction, elimination and dispersal [7,8]. In the proposed BFOS-tuning method, a cost function is defined in a systematic way such that the centers of the triangular membership functions and the PID gains for all fuzzy control rules can be selected as parameters to be determined. The performance of this controller will be verified by the simulation results.

2 Fuzzy PID Controllers

In fuzzy PID controllers, the input variables of the fuzzy rules are the error signals and their derivatives, while the output variables are the PID gains. The rules of a double-input and single input (DISO) fuzzy PID controller are usually expressed as

R_{ij} : IF x is A_i and y is B_j

$$\text{THEN} \quad u^{ij}(t) = K_P^{ij} e(t) + K_I^{ij} \int e(t)dt + K_D^{ij} \dot{e}(t) \tag{1}$$

where x denotes e(t) and $x \in X$, y denotes $\dot{e}(t)$ and $y \in Y$, i=1,2,...,n, j=1,2 ,...m, u(t) denotes output variable. Employing singleton fuzzifier, sum-product inference, and center-average defuzzifier, the output of a fuzzy PID controller is expressed as

$$u(t) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}(x,y)\, u^{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}(x,y)} \tag{2}$$

where $w_{ij}(x,y) = A_i(x)\, B_j(y)$ is the firing strength of the rule denoted as Rij.

The membership functions of input variables commonly adopt triangular functions in order to convenience calculation. In this paper, we assume that X and Y are universes of discourse for input variables e and $\dot{e}$ respectively. $\{A_i(x) \in F(X), i = 1,2,...,n\}$ are a cluster of fuzzy sets on X with triangular membership functions as shown in Fig. 1. The apexes of $\{A_i\}$ are denoted as xi, and conform to $x_1 < x_2 < ... < x_n$. The membership functions of $\{A_i\}$ can be calculated by

$$A_i(x) = \begin{cases} (x - x_{i-1})/(x_i - x_{i-1}), & x \in [x_{i-1}, x_i], \quad i = 2,3,...,n \\ (x_{i+1} - x)/(x_{i+1} - x_i), & x \in [x_i, x_{i+1}], \quad i = 1,2,...,n-1 \\ 1, & x < x_1 \text{ or } x > x_n \end{cases} \tag{3}$$

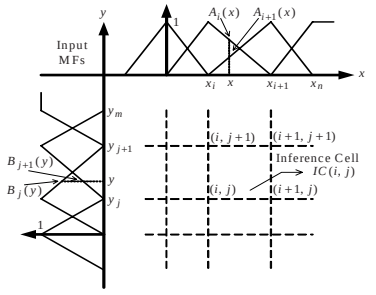


Fig. 1. The membership function of $\{A_i\}$ and $\{B_j\}$

Similar to $\{A_i\}$, $\{B_j(y) \in F(Y),\ j = 1,2,...,m\}$ are also a cluster of fuzzy sets on Y with triangular membership functions as shown in Fig. 1, the apexes of $\{B_j\}$ are denoted as yj, and conform to $y_1 < y_2 < ... < y_m$. The membership functions of $\{B_j\}$ can be calculated by

$$B_j(y) = \begin{cases} (y - y_{j-1})/(y_j - y_{j-1}), & y \in [y_{j-1}, y_j], \quad j = 2,3,...,m \\ (y_{j+1} - y)/(y_{j+1} - y_j), & y \in [y_j, y_{j+1}], \quad j = 1,2,...,m-1 \\ 1, & y < y_1 \ or \ y > y_m \end{cases} \tag{4}$$

The rule base plane can be decomposed into many inference cells (ICs) with output rules on its four corners, as shown in Fig. 1. The inference can be operated on these ICs. Assume that xi and xi+1 are any two adjacent apexes of $\{A_i\}$, yj and yj+1 are any two adjacent apexes of $\{B_j\}$. $x_i \le x \le x_{i+1}$, $y_j \le y \le y_{j+1}$ forms an inference cell IC(i,j) in $X \times Y$ input space. See Fig. 1. On the activated inference cell IC(i,j), the output of the fuzzy PID controller adopts dualistic piecewise interpolation functions of parameters of rule consequences, as follows [9]

$$\begin{aligned} u(t) = &[\sum_{s=i}^{i+1}\sum_{t=j}^{j+1} w_{st}(x,y)\, K_P^{st}]\, e(t) + [\sum_{s=i}^{i+1}\sum_{t=j}^{j+1} w_{st}(x,y)\, K_I^{st}] \int e(t)dt \\ &+ [\sum_{s=i}^{i+1}\sum_{t=j}^{j+1} w_{st}(x,y)\, K_D^{st}]\, \dot{e}(t) \end{aligned} \tag{5}$$

Equation (5) is an analytical model of fuzzy PID controller.

3 BFOS-Based Optimal Fuzzy PID Controller Design

The output trajectory of a sign-symmetry system is symmetrical to the original when the initial conditions and inputs are changed in sign. In many case, just like the AMB controller discussed in this paper, the system that has nonlinear control plant is also

sign-symmetry. We select that the input variables e(t) and $\dot{e}(t)$ are divided into 5 fuzzy sets named as Negative Big (NB), Negative Small (NS), Zero (ZO), Positive Small (PS), Positive Big (PB), respectively. The 5 fuzzy sets employ 50% overlapped triangular membership functions on the universe of discourse. Thus the parameters of input variables can he simplified to 4: the apex position x4 of PS and x5 of PB for e(t), y4 of PM and y5 of PB for $\dot{e}(t)$. The PID expressions of fuzzy control rule consequences, each has 3 modulus, are sign-symmetry. Therefore, when the membership functions of input variables are symmetrical to 0, there are only 15 independent rules in the 25 fuzzy control rules described in (1). In this way, there are totally 49 parameters to be determined when implementing a BFOS.

Natural selection tends to eliminate animals with poor foraging strategies and favor the propagation of genes of those animals that have successful foraging strategies, since they are more likely to enjoy reproductive success. After many generations, poor foraging strategies are either eliminated or shaped into good ones. This activity of foraging led the researchers to use it as an optimization process. The E. coli bacteria that are present in our intestines also undergo a foraging strategy. The control system of these bacteria that dictates how foraging should proceed can be subdivided into four sections: chemotaxis, swarming, reproduction, and elimination and dispersal:

(a) Chemotaxis: This process in the control system is achieved through swimming and tumbling via flagella. Therefore, an E. coli bacterium can move in two different ways: it can run (swim for a period of time) or it can tumble, and alternate between these two modes of operation in its entire lifetime. To represent a tumble, a unit length random direction, say, $\phi(j)$, is generated; this will be used to define the direction of movement after a tumble. In particular

$$\theta^i(j+1,k,l) = \theta^i(j,k,l) + C(i)\phi(j) \tag{6}$$

where $\theta^i(j,k,l)$ represents the ith bacterium at jth chemotactic, kth reproductive and lth elimination and dispersal step. C(i) is the size of the step taken in the random direction specified by the tumble (run length unit).

(b) Swarming: When a group of E. coli cells is placed at the centre of a semisolid agar with a single nutrient sensor, they move out from the centre in a traveling ring of cells by moving up the nutrient gradient created by consumption of the nutrient by the group. The spatial order results from outward movement of the ring and the local releases of the attractant; the cells provide an attraction signal to each other so they swarm together. The mathematical representation for swarming can be represented by:

$$\begin{aligned} J_{cc}(\theta, P(j,k,l)) &= \sum_{i=1}^{S} J_{cc}^{i}(\theta, \theta^i(j,k,l)) \\ &= \sum_{i=1}^{S}[-d_{att}\exp(-\omega_{att}\sum_{m=1}^{p}(\theta_m - \theta_m^i)^2)] \quad + \sum_{i=1}^{S}[-h_{rep}\exp(-\omega_{rep}\sum_{m=1}^{p}(\theta_m - \theta_m^i)^2)] \end{aligned} \tag{7}$$

where $J_{cc}(\theta, P(j,k,l))$ is the cost function value to be added to the actual cost function to be minimized to present a time varying cost function. S is the total number of bacteria, p the number of parameters to be optimized that are present in each

bacterium and d_{att}, ω_{att}, h_{rep}, ω_{rep} are different coefficients that are to be chosen properly.

(c) Reproduction: The least healthy bacteria die and the other healthiest bacteria each split into two bacteria, which are placed in the same location. This makes the population of bacteria constant.

(d) Elimination and dispersal: It is possible that in the local environment the life of a population of bacteria changes either gradually (e.g. via consumption of nutrients) or suddenly due to some other influence. Events can occur such that all the bacteria in a region are killed or a group is dispersed into a new part of the environment. They have the effect of possibly destroying the chemotactic progress, but they also have the effect of assisting in chemotaxis, since dispersal may place bacteria near good food sources. From a broad perspective, elimination and dispersal are parts of the population-level long-distance motile behavior. This Section is based on the work in [7]. As this paper concentrate on applying the new method to fuzzy PID design an in-depth discussion over the bacterial foraging strategy is not given here. The detailed mathematical derivations as well as theoretical aspect of this new concept are presented in [7]. The overall flowchart is shown in Fig. 2.

To evaluate the controller performance and get the satisfied transient dynamic, the cost function includes not only the four main transient performance indices, overshoot, rise time, settling time and cumulative error, but also the quadratic term of control input to avoid that the control energy became too big. The cost function is designed as

$$J = 1\Big/\int_0^\infty [\omega_1 te^2(t) + \omega_2 u^2(t)]dt + \omega_3 t_r + \omega_4 \sigma + \omega_5 t_s \tag{8}$$

where e(t) is the system error, u(t) is the controller input, t_r is the rise time, σ is the maximal overshoot, t_s is the settling time with 5% error band, $\omega_1, \omega_2, \omega_3, \omega_4, \omega_5$ are weighting coefficients. This research has picked the weighting coefficients $\omega_i = 0.2,\ i = 1,2,\ldots,5$ to cover all the performance indices completely.

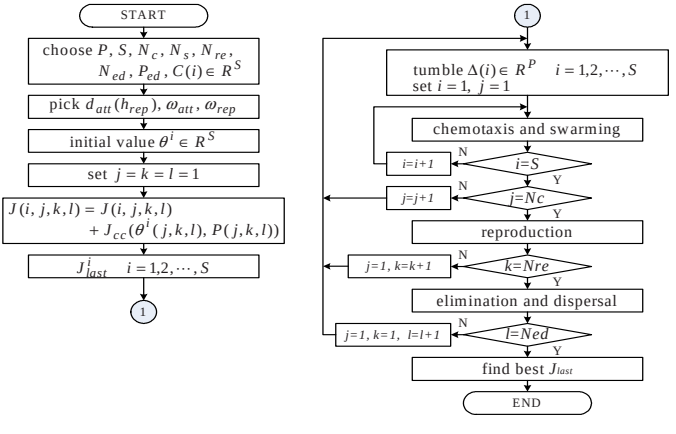


Fig. 2. The overall flowchart of BFOS

4 Simulations and Discussion

Fig. 3 shows the schematic of the controlled AMB system. It consists of a levitated object (rotor) and a pair of opposing E-shaped controlled-PM electromagnets with coil winding. An attraction force acts between each pair of hybrid magnet and extremity of the rotor. The attractive force each electromagnet exerts on the levitated object is proportional to the square of the current in each coil and is inversely dependent on the square of the gap. The entire system becomes only one degree of freedom of one axis, namely the axial position. Assuming a minimum distance to the length of the axis, the two attraction forces assure the restriction of radial motions of the axis in a stable way. The rotor position in axial direction is controlled by a closed loop control system, which is composed of a non-contact type gap sensor, a fuzzy PID controller and an electromagnetic actuator (power amplifier). This control is necessary since it is impossible to reach the equilibrium only by permanent magnets.

The rotor with mass m is suspended. Two attraction forces F1 and F2 are produced by the hybrid magnets. The applied voltage E from power amplifier to the coil will generate a current i which is necessary only when the system is subjected to an external disturbance w. Equations governing the dynamics of the system, linearized at the operation point (y=yo, i=0), are

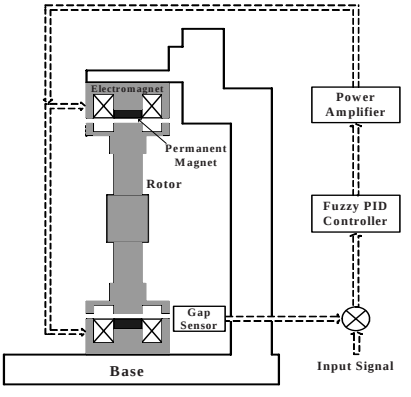


Fig. 3. The schematic of the controlled AMB system

$$\frac{d}{dt}\begin{bmatrix}\Delta y\\ \Delta \dot{y}\\ \Delta i\end{bmatrix}=\begin{bmatrix}0 & 1 & 0\\ a_{21} & 0 & a_{23}\\ 0 & a_{32} & a_{33}\end{bmatrix}\begin{bmatrix}\Delta y\\ \Delta \dot{y}\\ \Delta i\end{bmatrix}+\begin{bmatrix}0\\ 0\\ b\end{bmatrix}E+\begin{bmatrix}0\\ d\\ 0\end{bmatrix}w \tag{9}$$

where

$$a_{21}=\frac{1}{m}\frac{\partial \Delta F}{\partial y},\quad a_{23}=\frac{1}{m}\frac{\partial \Delta F}{\partial \Delta i},\quad a_{32}=-\frac{N}{L}\frac{\partial \Delta \phi}{\partial y} \tag{10}$$

$$a_{33} = -\frac{R}{L}, \quad b = \frac{1}{L}, \quad d = \frac{1}{m}, \quad L = N\frac{\partial \Delta\phi}{\partial \Delta i} \tag{11}$$

y is the distance from gap sensor to bottom of rotor. R and N are the resistance and number of turns of the coil. ϕ_1 and ϕ_2 are the flux of the top and bottom air gap, respectively.

To demonstrate the feasibility of the proposed scheme, the fuzzy PID controller is designed based on the proposed BFOS. The searched 49 optimal parameters including the triangular membership functions and the PID gains of fuzzy control rules are shown in Table 1 and Table 2, respectively.

Table 1. Center and width values of the optimized membership function ($\times 10^{-3}$)

	$e(t)$		$\dot{e}(t)$	
	center	width	center	width
NB	-2.3702	0.9235	-892.02	249.36
NS	-2.0765	2.3702	-750.64	892.02
ZO	0	4.1530	0	1501.28
PS	2.0765	2.3702	750.64	892.02
PB	2.3702	0.9235	892.02	249.36

Table 2. The optimized PID gains of fuzzy control rules

K_P^{ij}		$e(t)$				
		NB	NS	ZO	PS	PB
$\dot{e}(t)$	NB	3.2072	2.5176	5.8527	3.0434	3.3078
	NS	2.5176	2.3537	5.7214	3.9418	2.1030
	ZO	5.8527	5.7214	5.4519	3.3293	4.9865
	PS	3.0434	3.9418	3.9418	4.1751	3.8631
	PB	3.3078	2.1030	4.9865	3.8631	2.9493

K_I^{ij}		$e(t)$				
		NB	NS	ZO	PS	PB
$\dot{e}(t)$	NB	47.054	26.107	98.298	39.332	49.041
	NS	26.107	32.259	43.730	12.285	71.399
	ZO	98.298	43.730	20.220	81.691	4.5374
	PS	39.332	12.285	81.691	23.286	44.780
	PB	49.041	71.399	4.5374	44.780	21.041

K_D^{ij}		$e(t)$				
		NB	NS	ZO	PS	PB
$\dot{e}(t)$	NB	0.078019	0.035901	0.059289	0.037525	0.054979
	NS	0.035901	0.023338	0.066779	0.050989	0.090065
	ZO	0.059289	0.066779	0.061549	0.016826	0.099087
	PS	0.037525	0.050989	0.016826	0.044090	0.068130
	PB	0.054979	0.090065	0.099087	0.068103	0.084035

The step responses of rotor position from the gap sensor in the AMB system using the optimized fuzzy PID controller and the optimized conventional PID controller are shown in Fig. 3. It shows that the fuzzy PID controller has remarkably reduced the overshoot and settling time compared with the optimized conventional PID controller. The fuzzy PID controller has achieved good performances in both transient and steady state periods. Fig. 4 shows the converging patterns of the PID parameters. The PID gains are adaptive and the fuzzy PID controller has more flexibility and capability than the conventional ones.

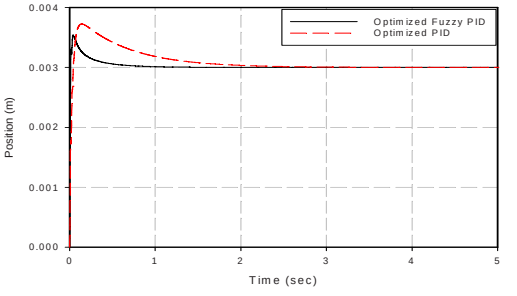


Fig. 4. The step responses of rotor position from the gap sensor in the AMB system using the optimized fuzzy PID controller and the optimized conventional PID controller

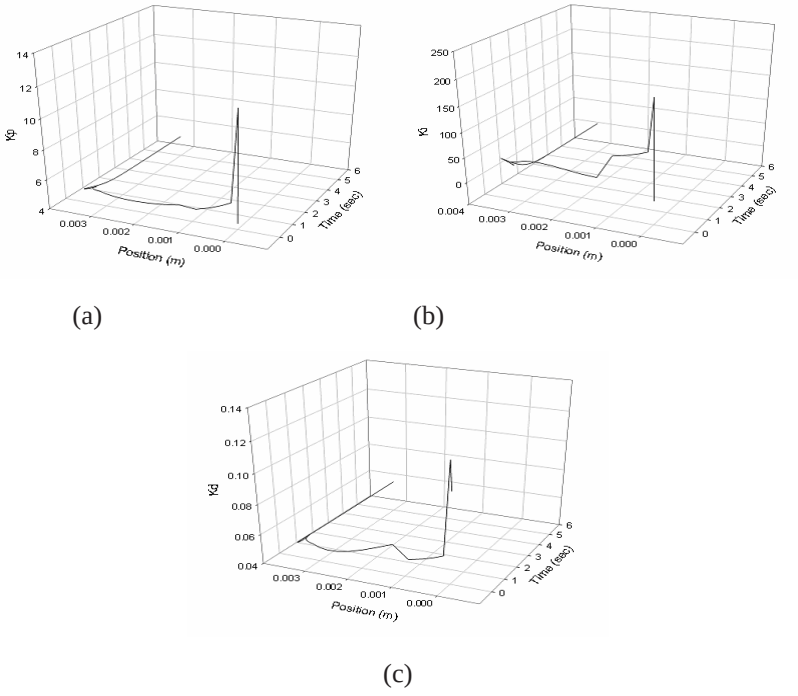


Fig. 5. The converging patterns of the PID parameters (a) Kp (b) Ki (c) Kd

5 Conclusions

This paper has proposed a bacterial foraging optimization scheme for the multi-objective optimization design of a fuzzy PID controller and applies it to the control of an AMB system. Another merit of the proposed scheme is the way to define the cost function based on the concept of multi-objective optimization. This scheme allows the systematic design of all major parameters of a fuzzy PID controller and then enhances the flexibility and capability of the PID controller. Since the PID gains generated by the proposed scheme are expressed in the form of fuzzy rules, they are more adaptive than the PID controller with fixed gains. The simulation results of this AMB system show that a fuzzy PID controller designed via the proposed BFOS has good performance.

Acknowledgments. The research was supported in part by the National Science Council of the Republic of China, under Grant No. NSC 96-2221-E-167-029-MY3.

References

1. Bennett, S.: Development of the PID Controller. IEEE Control Syst Mag. 13, 58–65 (1993)
2. Chen, G.: Conventional and Fuzzy PID Controller: An Overview. Int. J. Intell. Control Syst. 1, 235–246 (1996)
3. Na, M.G.: Auto-tuned PID Controller Using a Model Predictive Control Method for the Stream Generator Water Level. IEEE T Nucl. Sci. 48, 1664–1671 (2001)
4. Lin, L.L., Jan, H.Y., Shieh, N.C.: GA-based Multiobjective PID Control for a Linear Brushless DC Motor. IEEE T Mech. 8, 56–65 (2003)
5. Misir, D., Malki, H.A., Chen, G.: Design and Analysis of a Fuzzy Proportional- Integral-Derivative Controller. Int. J. Fuzzy Set Syst. 79, 297–314 (1996)
6. Tao, C.W., Taur, J.S.: Robust Fuzzy Control for a Plant with Fuzzy Linear Model. IEEE T Fuzzy Syst. 13, 30–41 (2005)
7. Passino, K.M.: Biomimicry for Optimization, Control, and Automation. Springer, London (2005)
8. Mishra, S.: A Hybrid Least Square-Fuzzy Bacterial Foraging Strategy for Harmonic Estimation. IEEE T Evolut. Comput. 9, 61–73 (2005)
9. Xiu, Z., Ren, G.: Optimization Design of TS-PID Fuzzy Controllers Based on Genetic Algorithms. In: 5th World Congress on Intelligent Control and Automation, Hangzhou, China, pp. 2476–2480 (2004)

An Analytical Adaptive Single-Neuron Compensation Control Law for Nonlinear Process

Li Jia[1,*], Peng-Ye Tao[1], Guang-Bo Chen[1], and Min-Sen Chiu[2]

[1] Shanghai Key Laboratory of Power Station Automation Technology, College of Machatronics Engineering and Automation, Shanghai University Shanghai 200072, China
[2] Department of Chemical and Biomolecular Engineering National University of Singapore Singapore, 119260
jiali@staff.shu.edu.cn

Abstract. To circumvent the drawbacks in nonlinear controller designing of chemical processes, an analytical adaptive single-neuron compensation control scheme is proposed in this paper. A class of nonlinear processes with modest nonlinearities is approximated by a composite model consisting a linear ARX model and a fuzzy neural network-based linearization error model. Motivated by the conventional feedforward control design technique in process industries, the output of FNNM can be viewed as measurable disturbance and a compensator can be designed to eliminate the disturbance influence. Simulation results show that the adaptive single-neuron compensation control plays a major role in improving the control performance, and the proposed adaptive control possesses better performance.

Keywords: Analytical, Single-neuron, Composite model, Compensator.

1 Introduction

The most widely used controllers in industrial chemical processes are based on conventional linear control technologies, such as PID controllers because of their simple structure and being best understood by process operators. However, demands for tighter environment regulation, better energy utilization, higher product quality and more production flexibility have made process operations more complex within a larger operating range. Consequently, conventional linear control technologies cannot give satisfactory control performance [1-3]. Recently, the focus has been shifting toward designing advanced control schemes based on nonlinear models. The nonlinear techniques using neural network and fuzzy system have received a great deal of attention for modeling chemical process, due to their inherent ability to approximate an arbitrary nonlinear function [4-9]. But several drawbacks arise in the controller designing based on nonlinear model: (1) it is time consuming procedure to rigorously develop and validate nonlinear models; (2) it is difficult to obtain the nonlinear model based control law in an analytical form; (3) the abundant linear controller designing methods and experience can not be easily used for nonlinear controller designing.

* Supported by Shanghai Leading Academic Discipline Project, Project Number: T0103 Correspondence should be addressed to JIA Li, associate professor.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 850–857, 2008.
© Springer-Verlag Berlin Heidelberg 2008

To circumvent above-mentioned problems, an adaptive single-neuron compensation control scheme is proposed in this paper. This paper is organized as follows. The proposed adaptive single-neuron compensation control is described clearly in Section 2. The performance analysis is discussed in Section 3. Simulation results are presented in Section 4 to illustrate the effectiveness of the proposed control scheme. Finally, concluding results are given in Section 5.

2 Adaptive Single-Neuron Compensation Control Scheme

In this Section, the proposed adaptive single-neuron compensation control will be described clearly and systematically.

2.1 Nonlinear Process Modeling

Fuzzy neural network (FNN) is a recently developed neural network-based fuzzy logic control and decision system, and is more suitable for online nonlinear systems identification and control. The FNN is a multilayer feedforward network, which integrates the TSK-type fuzzy logic system and RBF neural network into a connection structure. Considering of computational efficiency and easiness of adaptation, the zero-order TSK-type fuzzy rule is chosen in this paper, which is formulated as:

$$R^l : IF\ x_1\ is\ F_1^l\ AND\ x_2\ is\ F_2^l\ \cdots x_M\ is\ F_M^l,\quad THEN\ y = h_l \tag{1}$$

where R^l denote the lth fuzzy rule, N is the total number of fuzzy rules, $x \in R^M$ is the input variables of the system, $y \in R$ is the single output variable of the system. $F_i^l (l = 1, 2, \cdots, N, i = 1, 2, \cdots, M)$ are fuzzy sets defined on the corresponding universe [0, 1], and h_l is the lth weight or consequence of fuzzy system. The FNN contains five layers. The first layer is called input layer. The nodes in this layer just transmit the input variables to next layer. The second layer is composed of N groups (N represents the number of fuzzy IF-THEN rules), each of which is with M neurons (M represents the number of rule antecedent, namely the dimension of the input variables). Neurons of the lth group in the second layer receive inputs from every neurons of the first layer. And every neuron in lth group acts as a membership function. Here Gaussian membership function is chosen:

$$\exp(-\frac{(x_i - c_{li})^2}{\sigma_l^2}) \qquad i = 1, 2, \cdots, M; l = 1, 2, \cdots, N \tag{2}$$

The third layer consists of N neurons, which compute the fired strength of a rule. The lth neuron receives only inputs from neurons of lth group of second layer. There are two neurons in fourth layer. One of them connects with all neurons of the third layer through unity weight and another one connects with all neurons of the third

layer through h weights. The last layer has a single neuron to compute y. It is connected with two neurons of the fourth layer through unity weights.

The output of the whole FNN is:

$$y = \sum_{l=1}^{N} \Phi_l h_l \tag{3}$$

Where

$$\Phi_l = \frac{\exp(-\sum_{i=1}^{M} \frac{(x_i - c_{li})^2}{\sigma_l^2})}{\sum_{l=1}^{N} \exp(-\sum_{i=1}^{M} \frac{(x_i - c_{li})^2}{\sigma_l^2})}, (l = 1, 2, \cdots, N) \tag{4}$$

where the adjustable parameters of FNN are c_{li} , σ_l and h_l ($l = 1, 2, \cdots, N$; $i = 1, 2, \cdots, M$)

In this research, the FNN is employed to estimate the linearization error model of nonlinear process (denoted as FNNM) due to its capability of uniformly approximating any nonlinear function to any degree of accuracy. The SISO nonlinear process can be represented by following discrete nonlinear function

$$y(k+1) = f(\mathbf{x}(k)) \tag{5}$$

Where

$$\mathbf{x}(k) = (y(k), y(k-1), \cdots, y(k-n_y), u(k-n_d), u(k-n_d-1), \cdots, u(k-n_d-n_u))^T \tag{6}$$

where u and y are the input and output of the system, respectively.

The nonlinear model is

$$\hat{y}(k+1) = ARX(\mathbf{x}(k)) + FNNM(\mathbf{x}(k)) \tag{7}$$

In our research, Least Square identification method is used to identify ARX model. And the method proposed for the identification of FNNM can be found in our previous work [10].

2.2 Adaptive Single-Neuron Compensation Controller Design

Based on above-mentioned composite model, an adaptive single-neuron compensation control system is designed as Fig.1 and Fig.2, where G is the controlled nonlinear process, C represents the feedback controller, $\hat{G}_L$ represents the linear ARX model, FNNM is the error model and ASNCC represents the adaptive single-neuron compensation controller. The variables of r, y, y_L , e_L , u_f , respectively, represent the setpoint, the output of the controlled nonlinear process, the output of the ARX model, the error between y and the output of the ARX model and the compensation input.

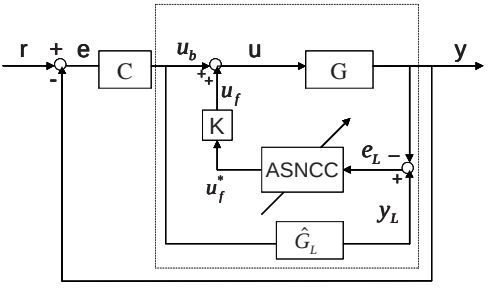


Fig. 1. The structure of Adaptive Single-Neuron Compensation control system

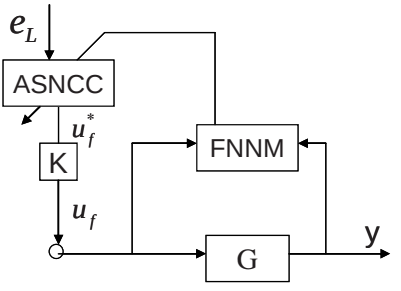


Fig. 2. The learning method for compensation controller

Inspired by the simple and widely used PID structure, a single neuron is adopted in the ASNCC controller structure. It has three inputs: $e_L(k)$ is the error between process output and the output of ARX model, $\Delta e_L(k) = e_L(k) - e_L(k-1)$ is the difference between the current and previous error, $\delta e_L(k) = \Delta e_L(k) - \Delta e_L(k-1) = e_L(k) - 2e_L(k-1) + e_L(k-2)$. And $w_i (i = 1,2,3)$ are the neuron weights.

The control law of the ASNC controller is obtained as following:

$$\begin{aligned} u_f^*(k) &= u_f^*(k-1) + \Delta u_f^*(k) \\ \Delta u_f^*(k) &= w_1(k)e_L(k) + w_2(k)\Delta e_L(k) + w_3(k)\delta e_L(k) \end{aligned} \tag{8}$$

The response of y can be obtained as

$$y = Gu = Gu_b + Gu_f \ . \tag{9}$$

Where

$$u_f = ASNCC(e_L) \ . \tag{10}$$

We seek to determine a $ASNCC(e_L \mid W)$ such that the criterion E is minimized with response to the set of parameter quadratic matrix W

$$E(k)=\frac{1}{2}(y_L(k)-y(k))^2+\lambda(u(k-1)-u(k-2)) \ . \tag{11}$$

where $y_L=\hat{G}_L u_b$.

If ASNCC has completely learned the error characteristics, ie $E(k)=0$, we have

$$y_L=y \ . \tag{12}$$

Combining the result with that of equation (18) yields

$$\hat{G}_L u_b=Gu_b+Gu_f=Gu_b+G\ ASNCC(e_L) \ . \tag{13}$$

Then we can get

$$ASNCC(e_L)=-G^{-1}(G-\hat{G}_L)u_b \ . \tag{14}$$

Since $u_b=C(r-y)$, y can also be written as

$$y=(1+\hat{G}_L C)^{-1}\hat{G}_L Cr \ . \tag{15}$$

Obviously, the closed-loop performance is improved because the nonlinear process to be controlled by linear controller is linearized. As a result, conventional linear control technologies, such as PID control can be directly applied. Next Section we will introduce how to find optimal W for ASNCC.

2.3 The Algorithm for Adaptive Single-Neuron Compensation Control System

In the following, a learning algorithm will be developed to update the parameters of the ASNCC controller at each sampling instant.

Consider the object function $E(k)=\frac{1}{2}(y_L(k)-y(k))^2$, the control object is to find the optimal values of w_i to minimize the object function.

Since the controller parameters w_i (i=1,2,3) are constrained to be positive or negative, the following mapping function is introduced:

$$w_i(k)=\begin{cases} e^{\varsigma_i(k)} & (w_i(k)\geq 0) \\ -e^{\varsigma_i(k)} & (w_i(k)<0) \end{cases} \quad i=1,2,3 \ . \tag{16}$$

where ς_i is a real number.

To tune the controller parameters at every sampling time, backpropogation method is used to derive the parameter updating equations as follows:

$$\begin{aligned} \varsigma_i(k+1) &= \varsigma_i(k)-\eta\frac{\partial E}{\partial \varsigma_i(k)} \\ &= \varsigma_i(k)-\eta\frac{\partial E}{\partial w_i(k)}w_i(k) \end{aligned} \ . \tag{17}$$

$$\frac{\partial E}{\partial w_i(k)} = \frac{\partial E}{\partial y(k)} \frac{\partial y(k)}{\partial u(k-1)} \frac{\partial u(k-1)}{\partial w_i(k)} = -(y_L(k) - y(k)) \frac{\partial y(k)}{\partial u(k-1)} \frac{\partial u(k-1)}{\partial w_i(k)} \quad (18)$$

for i=1~3, $\frac{\partial u(k)}{\partial w_1(k)} = e_L(k)$, $\frac{\partial u(k)}{\partial w_2(k)} = \Delta e_L(k)$, $\frac{\partial u(k)}{\partial w_3(k)} = \delta e_L(k)$, and $\frac{\partial y(k)}{\partial u(k-1)}$ can be obtained from the FNMM and the ARX model.

3 Stability Analysis

For the feedback controller C, the ASNCC now becomes a new nonlinear uncertainty adding to the loop. Designing the controller to be a robust one, we need to account for the uncertain factors due to modelling uncertainties and ASNCC.
Suppose the following:

Assumption 1: The modelling nonlinear uncertainty ΔG satisfies

$$\Delta G(0,t) = 0, \| \Delta G(\mathrm{u},t) \|_2 \le \gamma \| u \|_2, \gamma < \infty$$

Assumption 2: The "disturbance" e_L is bounded by $\|e_L\| \le \alpha$.

Assumption 3: The nominal subsystem Ω_1 and Ω_2 are stable, where

$$\Omega_1 = \frac{C}{(1+G_L C)}$$

$$\Omega_2 = \frac{G_L C}{(1+G_L C)}$$

Assumption 4: If ASNCC is trained properly, the output of ASNCC must be bounded, i.e.

$$\left\| ASNCC(e_L) \right\|_2 \le \varphi \left\| e_L \right\|_2 \qquad \varphi < \infty$$

Under above-mentioned assumptions, we have the following theorem:
Theorem: Due to the limited pages, the proof is deleted.

4 Example

The example considered is the van de Vusse reaction with the following reaction kinetic scheme: A→B→C,A→D, which is carried out in an isothermal CSTR. The dynamics of the reactor are described by the following equations [11-12].

$$\begin{aligned} \frac{dC_A}{dt} &= -k_1 C_A - k_3 C_A^2 + \frac{F}{V}(C_{Af} - C_A) \\ \frac{dC_B}{dt} &= k_1 C_A - k_2 C_B - \frac{F}{V} C_B \end{aligned} \quad (19)$$

where the parameters used are: $k_1 = 50\,\mathrm{h}^{-1}$, $k_2 = 100\,\mathrm{h}^{-1}$, $k_3 = 10\mathrm{L}/(\mathrm{mol}\cdot\mathrm{h})$, $C_{Af} = 10\,\mathrm{mol/L}$, and $V = 1\mathrm{L}$. The nominal operation condition is $C_A = 3.0\,\mathrm{mol/L}$, $C_B = 1.12\,\mathrm{mol/L}$, and $F = 34.3\mathrm{L/h}$. The control objective is to regulate the concentration of component B (C_B) by manipulating the inlet flow rate F.

Independent random signal with uniform distribution between [20, 45] is used to simulate 100 input-output data points. To proceed with the proposed algorithm, $\mathbf{x}(k)$ is chosen to be $\mathbf{x}(k) = [C_B(k), C_B(k-1), F(k)]^T$, Then generated input-output data points are used to identify the proposed composite model, i.e. a linear ARX model and a FNNM model. Based on the obtained composite model, the adaptive single-neuron compensation control can be designed next.

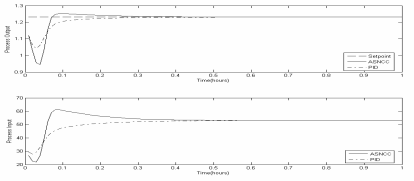


Fig. 3. The comparison of the results between ASNCC and PID controller (I)

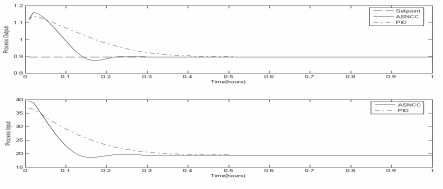


Fig. 4. The comparison of the results between ASNCC and PID controller (II)

The initial $\varsigma_1 = -2$, $\varsigma_2 = 0.8$, $\varsigma_3 = -2$, $w_1 = e^{-2}$, $w_2 = e^{0.8}$, $w_3 = e^{-2}$, $K = 0.01$, the learning rate $\eta = 0.3$ is set for ASNC controller. The tuning parameters of feedback controller C is set based on IMC based PID tuning rule.

To evaluate the control performance of the proposed controller and PID controller, 10% and -20% step changes in the setpoint of C_B are considered. Fig.3 and Fig.4 show that the response of proposed adaptive single-neuron compensation control system can get better performance than the PID control system.

5 Conclusion

In this paper, an adaptive single-neuron control scheme is proposed. A class of nonlinear processes with modest nonlinearities is approximated by a composite model consisting a linear ARX model and a fuzzy neural network-based linearization error

model. Simulation results show that the adaptive single-neuron compensation control plays a major role in improving the control performance, and the proposed adaptive control possesses better performance.

References

1. Kim, S.J., Lee, M., Park, S.: A Neural Linearizing Control Scheme for Nonlinear Chemical Processes. Computer Chem. Engng. 21, 187–200 (1997)
2. Gao, F., Wang, F.L., Li, M.Z.: An Analytical Predictive Control Law for a Class of Nonlinear Process. Ind. Eng. Chem. Res. 39, 2029–2034 (2000)
3. Zhang, J.: A Nonlinear Gain Scheduling Control Strategy Based on Neuro-fuzzy Networks. Ind. Eng. Chem. Res. 40, 3164–3170 (2001)
4. Aswin, N.V., Ravindra, D.G.: Fuzzy Segregation-based Identification and Control of Nonlinear Dynamic System. Ind. Eng. Che. Res. 41, 538–552 (2002)
5. Cao, S.G., Rees, N.W., Feng, G.: Analysis and Design for a Class of Complex Control Systems Part 1: Fuzzy Modeling and Identification. Automatica 33, 1017–1028 (1997)
6. Narendra, K.S.K., Parthasarathy: Identification and Control of Dynamica Systems using Neural Networks. IEEE Trans. Neural Networks 1, 4–27 (1990)
7. Sugeno, M., Tanaka, K.: Successive Identification of a Fuzzy Model and Its Application to Prediction of A Complex System. Fuzzy Sets and Systems 42, 315–334 (1991)
8. Wang, L., Langari, R.: Complex System Modeling via Fuzzy Logic. IEEE Trans. System, Man and Cybernetics-Part B: Cybernetics 26, 100–106 (1996)
9. Wang, Y., Rong, G.: A Self-organizing Neural-network-based Fuzzy System. Fuzzy Sets and Systems 103, 1–11 (1999)
10. Li, J., Tao, P.Y., Chiu, M.S.: Fuzzy Neural Network-Based Adaptive Single Neuron Controller. In: Li, K., Fei, M., Irwin, G.W., Ma, S. (eds.) LSMS 2007. LNCS, vol. 4688, pp. 412–423. Springer, Heidelberg (2007)
11. Dolye, F.J., Ogunnaike, A., Pearson, R.K.: Nonlinear Model-based Control Using Second-order Volterra Models. Automatic 31, 697–714 (1995)
12. Eskinat, E., Johoson, E.H., Luyben, W.L.: Use of Hammerstein Models in Identification of Nonlinear Systems. AIChE J. 37, 255–268 (1991)

Energy-Saving Control System of Beam-Pumping Unit Based on Radial Basic Function Network

Jing-Wen Tian[1,2], Mei-Juan Gao[1,2], Shi-Ru Zhou[1], and Fan Zhang[2]

[1] Beijing Union University, Department of Automatic Control, Postfach 10 01 01, Beijing, China
[2] Beijing University of Chemical Technology, School of Information Science, Postfach 10 00 29, Beijing, China
ttjjww999@163.com, ttgg500@163.com

Abstract. Considering the issues that the energy saving process for beam-pumping unit is a complicated and nonlinear system, and it is very difficult to found the process model to describe it. The radial basic function network (RBFNN) has the ability of strong nonlinear function approach and adaptive learning and also has the feature of fast convergence. In this paper, an intelligent energy-saving control system of beam-pumping unit based on RBFNN is presented. We construct the structure of RBFNN, and adopt the K-Nearest Neighbor algorithm and least square method to train the network. The parameters of energy-saving control process of beam-pumping unit are measured using multi sensors, and then the system can control the working state of beam-pumping unit real-time. The system is used in the oil recovery plant. The experimental results prove that this system is feasible and effective.

Keywords: Beam-pumping unit, Energy-saving, Control system, Radial basic function network.

1 Introduction

The low durchgriff reservoir has low generating oil capability. Especially the gushing well is little while the oil recovery is into middle and end period. The method of water displacing oil is adopted for most oil field. When the oil is displaced to oil well, the beam-pumping units are used to lift the oil from the well up to ground. The most power consume of oil production are contributed by the water displacing oil process and the electrical energy consumption of beam-pumping units. So the research for energy saving of beam-pumping units is very important and necessary [1]-[2].

In a general way, the power efficiency of beam-pumping units is selected according to the maximal pumping quantity of oil well and also some design margin is remained under most circumstance. But along with the oil producing continuously, the oil production decreases gradually, and the liquid level in oil well decrease low gradually. So the oil liquid in pump is very low fullness, the beam-pump unit is running at low efficiency, even sometime the pump run in empty condition, if the low efficiency and empty pump running are not controlled, then lots of electric energy will be wasted.

For the low permeability and low off-take potential oil field, the effective method of beam-pumping unit energy-saving is to dynamic control the stop time of beam-pump unit

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 858 – 865, 2008.
© Springer-Verlag Berlin Heidelberg 2008

to attain the energy-saving. Traditional controlling for the stop time of beam-pump unit is direct turnon and stop by the experienced workers or time turnon and stop by manual work [3], and the work frequency of beam-pumping unit running is not adjusted. These method can achieve energy-saving at a certain extent, but there are some problems exist in them such as affection the oil production from ground acquisition due to the stop time longer or the liquid fullness of oil pump underground lower and empty pump due to the stop time shorter, thereby the waste energy is still exist at a certain extent. So traditional method cannot solve the problem that the work capability of beam-pumping unit dynamic response the change of oil well load, and cannot achieve the aim of energy-saving better, and also can affect the oil production.

The ability to approach nonlinear function of BP network has been proved in theory also has been validated in actual applications [4]. BP network is one kind of global approach network, the every weight value of the network is adjusted to every sample while the network is trained, this conduce the speed of network learning slowly. On the other hand, the BP algorithm may converge to a local minimum. The radial basic function network adopted local approaching network. It has a superiority of fast learning speed and an ability of strong non-linear function approach and mode-classification [5]. Meanwhile the radial basic function network has the simple implementation process. So the radial basic function network was applied to the intelligent energy-saving control of beam-pumping unit in this paper.

2 Method Study

2.1 Radial Basic Function Neural Network (RBFNN)

Radial basic function network is constructed of three layers. Those are input layer, hidden layer and output layer. The structure of RBFNN is shown in Fig. 1.

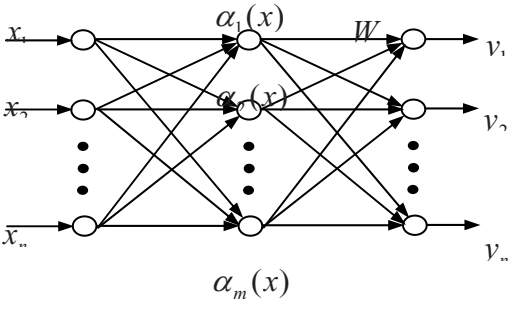


Fig. 1. The structure of radial basic function network

The input layer nodes transmit input signal to reach hidden layer, hidden layers nodes are described by the Gauss kernel function and output layer nodes are described by the linear function. The kernel function of hidden layer nodes responds to the input signal in local. When the input signal close the centre range of the Gauss kernel, the hidden layer nodes will produce larger output. So the radial basic function network is

local approaching network and it has a superiority of fast learning speed. In this paper the basic function is defined as follows:

$$\alpha_i(x) = \exp\left[\frac{-\| X - C_i \|^2}{2\sigma_i^2}\right] \quad i = 1,2,\cdots m \tag{1}$$

Here, $\alpha_i(x)$ is the output of number i node of hidden layer; $X = (x_1, x_2, \cdots, x_n)^T$ is the input samples; C_i is the centre of the Gauss kernel function, and it has the same dimensions with X; σ_i is the variable of number i of hidden layer, it is called standardized constant; m is the total number of the hidden layer nodes.

Input layer actualize the nonlinear mapping that is $X \to \alpha_i(x)$, output layer actualize the linear mapping $\alpha_i(x) \to y_k$, namely, the output of radial basic function network is the linear combination of hidden layer nodes outputs, the output of radial basic function network is given by follow formula:

$$y_k = \sum_{i=1}^{m} \omega_{ik} \alpha_i(x) \quad k = 1,2,\cdots, p \tag{2}$$

Here, ω_{ik} is the weight value of the network; p is the number of the output layer nodes.

2.2 K-NN Algorithm

K-NN algorithm (K-Nearest Neighbor) is one kind of pattern clustering algorithm, suppose N kinds of patterns $(C_1, C_2, \cdots, C_N)$, the typical sample of each kind pattern written as $T_n(n = 1,2,\cdots,n)$, to the unknown samples X, K-NN algorithm will decide its adscription kind by follow steps.

Suppose $d(x,y)$, $(x,y) \in V$, express the similarity degree of any two points x, y of space V, namely, the $d(x,y)$ is more small, the x, y is more similar, d can be considered some kind "distance". To any unknown sample X, the k typical samples $\{T_1, T_2, \cdots, T_k\}$ that have the least distance with the X can be obtained using $d(X, T_i)$, then we can calculate the number $\upsilon(n)$ that $T_i(i = 1,2,\cdots,k)$ belong to the pattern C_n. Finally, according to the maximal $\upsilon(n)$ to decide X pattern kind. That expressed as follow:

$$\text{if } \upsilon(m) = \max\{\upsilon(1), \upsilon(2), \cdots, \upsilon(n)\}$$

$$\text{then } X \in C_m$$

In practical application, k is used as 1, after the numbers of hidden layer nodes are ascertained, we can obtain the parameters of kernel function of each hidden layer nodes with the samples of each pattern clustering, the formula is as follows:

$$C_i = \frac{1}{M_i} \sum_{x \in \theta_i} X \quad i = 1,2,\cdots, m \tag{3}$$

$$\sigma_i^2 = \frac{1}{M_i} \sum_{x \in \theta_i} (X - C_i)^T (X - C_i) \quad i = 1,2,\cdots m \tag{4}$$

Here, θ_i represent the all samples of the i pattern clustering; M_i is the numbers of sample of the i pattern clustering.

3 Energy-Saving Control System

3.1 Beam-Pumping Unit Energy-Saving Affect Factors

The beam-pumping unit includes two parts which are the ground part and the underground part, the ground device include the dynamic force and balance of beam-pumping unit which composed of the pumping rod strain, well output liquid quantity, beam-pump unit current, beam-pump unit voltage and active power. The underground part include oil pump and corresponding valve, this is shown in Fig. 2.

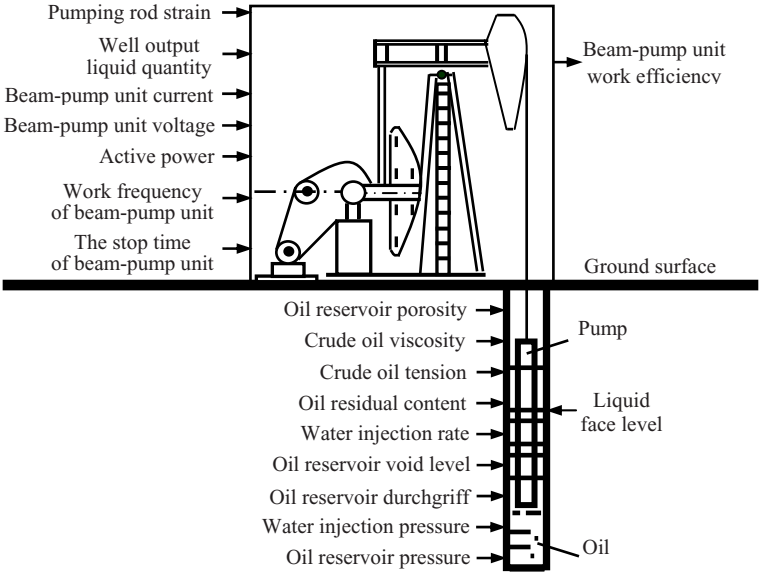


Fig. 2. The schematic diagram of beam-pumping unit with affect factors

The energy-saving problem of beam-pumping unit is to enhance the work efficiency up to the hilt, the work efficiency is expressed the output oil (kilogram)/ consumed electrical energy (kilowatt-hour). The work efficiency is decided by the working statue of the beam-pump units. Now the control for the stroke of beam-pump unit is difficult, so we make the stroke constant temporarily, then the controllable working statues of beam-pump units include two parts, one is the stop time of beam-pump unit, another is the beam-pump unit work frequency. The stop time of beam-pump unit and the beam-pump unit work frequency is affected by the oil output

velocity in oil well, the liquid level in the oil well and the liquid fullness of oil pump underground.

The oil well is several kilometers depth underground and the space of oil well is much small, it is very difficult to measure the liquid level in the oil well and the liquid fullness of oil pump underground directly. We can estimate the pump fullness indirectly by measuring the pumping rod strain, well output liquid quantity, beam-pump unit current, beam-pump unit voltage and active power. To measure the oil produce rate in well directly is difficult also, according to status quo of oil production, the 8 primal parameters are discovered which influence the oil produce rate. They are oil reservoir porosity, oil reservoir durchgriff, reservoir crude oil residual content, oil reservoir pressure, crude oil viscosity, crude oil tension, water injection pressure, water injection rate. So the working state of the beam-pumping units can be controlled according to the oil output rate in oil well, the liquid level in the oil well and the liquid fullness of oil pump underground, the work efficiency of beam-pumping can be enhanced and the electric energy of pump will be saved.

3.2 Energy-Saving Control System Architecture

According to the above describe, we know that the energy-saving problem for beam-pumping unit is very complicated nonlinear system in the water injection oil field, the classical control method can not solve this problem. So the radial basic function network internal model control project is adopted to solve the beam-pump unit energy-saving problem [6]. The control model is shown in Fig. 3.

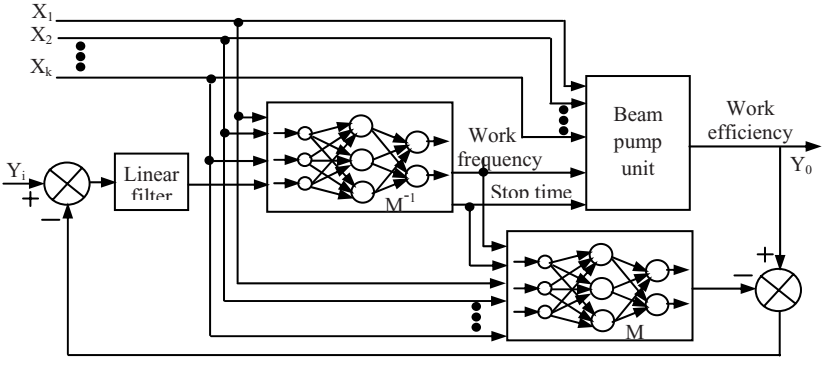


Fig. 3. The beam-pump unit energy saving control model

The energy-saving control system contains two radial basic function network M^{-1} and M with the same structure. The network M^{-1} is used as the controller of the system. The network M is used as the state estimator of the controlled object. The state estimator and controlled object are located in paralleled state, their difference is transmitted to input by the feedback way, then it is go through the linear filter into the controller M^{-1}, the output of M^{-1} is the input of the controlled object. Taking the state estimator M as the training object, the controller M^{-1} will learn the converse

dynamic behavior of controlled object indirectly by repeated training time after time, and the error of control system will approach to zero. The input of M is the input of the controlled object and the output of M is the output response of the controlled object.

The controller M^{-1} and state estimator M are composed of an input layer, on hidden layer and an output layer. The input layer of state estimator M has 7 neurons that are the stop time of the beam-pump unit, the beam-pump unit work frequency, oil reservoir porosity, crude oil viscosity, water injection pressure, water injection rate and the pumping rod strain. The output layer of state estimator M has 2 neurons that are the well output liquid quantity and beam-pump unit current, the ratio (well output liquid quantity /beam-pump unit current) is the work efficiency of beam-pumping unit. The input layer of state estimator M^{-1} has 7 neurons that are the oil reservoir porosity, crude oil viscosity, water injection pressure, water injection rate and the pumping rod strain, the feedback quantity of well output liquid quantity which is the error of system output and anticipate output and the feedback quantity of beam-pump unit current,. The output layer of state estimator M^{-1} has 2 neurons that are the stop time of the beam-pump unit and the beam-pump unit work frequency which is the controlled variable of controlled object.

In this control system, first the state estimator M is trained off-line using all known sample and then get the in-out property of the controlled object. The network weights of the trained M can be taken as the initial network weights of the controller M^{-1}. Then the M^{-1} is trained using the back-propagation error of the state estimator M, namely using the error between the anticipant output and the real output of control system to adjust the every network weights of M^{-1} to make the real output of system approach the anticipant output. The controller M^{-1} can be trained on-line and revised while the control system is running.

3.3 Control Process

In the working process of the control system, the pumping rod strain, well output liquid quantity, beam-pump unit current etc parameters were connected to corresponding sensors. The controlled variable stop time of the beam-pump unit and the beam-pump unit work frequency neurons of the input layer were connected to 2 controllers respectively. If the pumping rod strain, well output liquid quantity, beam-pump unit current etc was changing, the related sensors detected the change timely and inputted them to M^{-1} network, which could compute the control value of the stop time of the beam-pump unit and the beam-pump unit work frequency through the network compute, and further compute the work efficiency of beam-pumping unit through the network M. After this, if the work efficiency of beam-pumping unit also could not satisfy the anticipate effect, the M^{-1} network could be trained on-line using the error between the output of M^{-1} and the anticipant output until get the anticipate effect. That is mean to control the stop time of the beam-pumping unit and the beam-pumping unit work frequency real-time through several sensors [7]-[8].

4 Application

Before the system work, we collect the training samples. We can obtain large numbers of relations between the stop time of the beam-pump unit, the beam-pump unit work frequency, oil reservoir porosity, crude oil viscosity, water injection pressure, water injection rate and the pumping rod strain and the well output liquid quantity and beam-pump unit current by experiment, and then we can get many group training samples, and every training sample is composed of 7-input and 2-output.

In the working process, first, we train the controlled object state estimator M off-line. Then we use the trained M to train the feed-forward controller M^{-1}. After M^{-1} is confirmed, the system can work on-line and also the M^{-1} can be trained on-line. In the system control process, only the feed-forward controller M^{-1} takes effect.

We write the corresponding control program using C++. First, all training samples are clustering by the K-NN algorithm. According to the actual sample data and the program computing, we get 6 clustering. Then the hidden layer of RBFNN has 6 neurons also. The topology structure of the RBFNN is 7-6-2. In order to show the advantage and feasibility of RBFNN, we adopt RBFNN and BP network to control the stop time of the beam-pumping unit and the beam-pumping unit work frequency real-time. The structure of BP is the same as RBFNN. The learning parameters of network are selected as follow: the system maximum error is 0.01, the maximum error of single sample is 0.001, the iteration times of network is 5000.

We could random separate the 500 samples into 5 groups, and take out 70 sample data as training samples each time, while the other 30 samples as testing sample. After the normalized process of the sample set, the RBFNN and BP network was trained 5 times respectively using 5 groups different training sample, thenceforth we used the corresponding testing samples to test. The training error and testing error were respectively the average value of the 5 times training error and 5 times testing error. The result of train and test by RBFNN and BP network is shown in Table 1.

Table 1. Contrast of training and testing results by RBFNN and BP network

Method	Training error	Testing error	Average iteration times
Radial basic function network	0.0052	0.103	531
BP network	0.0089	0.237	1658

From the table 1, we can see the mean squared error of training samples of RBFNN is smaller than that of BP network. Moreover the mean squared error of testing samples of RBFNN is also smaller than that of BP network. Furthermore the iteration times of the RBFNN is obvious smaller than that of BP network with the same system error. It shows that the RBFNN is all superior to the BP network at the at the estimation accuracy, simulation accuracy and convergence rate aspects. So the control system based on the RBFNN has higher stability and faster real-time controlled speed.

5 Conclusion

Because the radial basic function network is local approaching network, it has a superiority of fast learning speed and an ability of strong function approach and mode-classification in relation to the BP network. Therefore the intelligent energy-saving control system of beam-pumping unit based on the radial basic function network in this paper can enhance control accuracy and control speed to a great extent. The control system is tested by experiment data, and the testing precision of the system is verified through the verification method of circular sample. The results show that the system can control the energy-saving process of beam-pumping unit real-time and effectively and conveniently.

References

1. Su, D.S., Liu, X.G., Lu, W.X., et al.: Overview of Energy-saving Mechanism of Beam-balanced Pump. Oil Machinery 19(5), 49–63 (2001)
2. Bai, L.P., Ma, W.Z., Yang, Y., et al.: The Discussing of Energy-saving of Motor for Beam-balanced Pump. Oil Machinery 27(3), 41–44 (1999)
3. Xu, F.R., Zhao, X.SH.: Summarization on Energy Saving of Electric Control Device for Oil Pumping Jack. Electric Drive Automation 26(5), 1–8 (2004)
4. Jiao, L.C.: Neural Network System Theory. Xi an Electronic Science and Technology University Press, Xian (1995)
5. Zhao, ZH.N., Xu, Y.M.: Introduction to Fuzzy Theory and Neural Networks and Their Application. Tsinghua University Press, Beijing (1996)
6. Li, S.Y.: Fuzzy Control Neuro-control and Intelligent Cybernetics. Harbin Institute of Technology Publishing House, Harbin (1996)
7. Ding, B., Zhang, Y.L., Sun, L.F.: RL Fuzzy Neural Network and its Application to Oil Pumping Control. Control Engineering of China 9(6), 57–59 (2002)
8. Qi, W.G., Zhu, X.L., Zhang, Y.L.: A Study of Fuzzy Neural Network Control of Energy-saving of Oil Pump. In: Proceeding of the CSEE, pp. 137–140 (2004)

ANN Combined-Inversion Control for the Excitation System of Generator*

Wan-Cheng Wang

College of Electrical Engineering, Hohai University, Nanjing, 210098
wcwang@hhu.edu.cn

Abstract. The implementation of the inversion based linearizing method for the excitation system usually requires the full state feedback. This severely restricts its practical use because some of these states are very difficult to measure by physical sensors in practice due to no sensor available in the market or its high cost. To address this issue, the combined-inversion is proposed by coupling the left-inversion used as a soft-sensor with the right-inversion used as a nonlinear controller. Furthermore, to overcome the difficulty in explicit implementation of the combined-inversion by analytic means, an artificial neural network (ANN) is used to approximate the combined-inversion, thus forming the ANN combined-inversion. The simulation results demonstrate the effectiveness of the ANN combined-inversion.

Keywords: Inversion, ANN, Excitation System.

1 Introduction

The excitation control is all along an important problem in power system control field, and it is mainly used to keep the generator terminal voltage stable. At the early stage of power system control, the excitation controller is just the simple the automatic voltage regulator (AVR)[1], [2]. With the expansion of the power system, the power system stabilizer (PSS) is developed [3], [4]. With the development of the nonlinear control method, the feedback linearizing method based on inversion has received a great deal of attention in recent years [5]. The implementation of the inversion based linearizing method for the excitation system usually requires the full state feedback. Unfortunately, due to no sensor available in the market or its high cost, some of these states are very difficult to measure by physical sensors in practice, hence severely restricting its practical use.

To address this issue, a new concept of combined-inversion is proposed by coupling the left-inversion with the right-inversion. Then, the ANN combined-inversion is obtained by using a static ANN to approximate the combined-inversion so as to overcome the difficulty in explicit implementation of the combined-inversion by analytic means. The ANN combined-inversion is capable of linearizing the excitation system. For the linearized system, a traditional PID controller is used to implement

* This work is supported by the Natural Science Foundation of Hohai University (2007418111).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 866–873, 2008.
© Springer-Verlag Berlin Heidelberg 2008

the ultimate closed control of the excitation system. The simulation results show its validity.

2 The Combined-Inversion of the Excitation System

Consider the excitation system of the single-machine infinite-bus power system with double transmission lines as shown in Fig. 1, which can be expressed as follows [6]

$$\begin{cases} \dot{\delta} = \omega - \omega_0 \\ \dot{\omega} = \dfrac{\omega_0}{H} P_{m0} - \dfrac{D}{H}(\omega - \omega_0) - \dfrac{\omega_0 E_q' V_s}{H x_{d\Sigma}'} \sin\delta - \dfrac{\omega_0 V_s^2}{2H}\left(\dfrac{x_d' - x_q}{x_{d\Sigma}' x_{q\Sigma}}\right) \sin 2\delta \\ \dot{E}_q' = -\dfrac{x_{d\Sigma}}{x_{d\Sigma}' T_{d0}} E_q' + \dfrac{(x_d - x_d')V_s \cos\delta}{x_{d\Sigma}' T_{d0}} + \dfrac{U_{E_q'}}{T_{d0}} \end{cases} \tag{1}$$

where $\left(\delta, \omega, E_q'\right)^T$ is the vector of states, the excitation voltage is the control input $u = U_{E_q'}$, and controlled output is the generator terminal voltage

$$y = V_t = \sqrt{\left(\frac{x_q V_s \sin\delta}{x_{q\Sigma}}\right)^2 + \left(\frac{E_q'\left(x_{d\Sigma}' - x_d'\right) + V_s x_d' \cos\delta}{x_{d\Sigma}'}\right)^2} = y\left(E_q', \delta\right) . \tag{2}$$

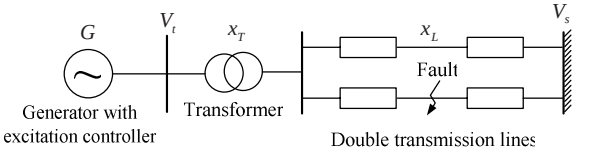


Fig. 1. Single machine infinite-bus system

It should be noted that the variables ω is directly measurable and while the $\delta, E_q^{'}$ are directly immeasurable. The meaning of all the above symbols is as follows: δ: the power angle of the generator; ω: the relative speed of the generator; $E_q^{'}$: the Electromotive Force (EMF) in the q axis of generator; $U_{E_q'}$: the excitation voltage; V_t: the generator terminal voltage; D: the per unit damping constant; H: the per unit inertia constant; ω_0: the synchronous speed, $\omega_0 = 2\pi f$; P_{m0}: initial static value of total mechanic power; T_{d0}: the d axis transient open circuit time constant; x_d: the d axis reactance of generator; x_q: the q axis reactance; x_d': the d axis transient reactance; x_T: the reactance of the transformer; x_L: the reactance of the transmission line;

V_s : the infinite bus voltage (normally $V_s = 1$); $x_{d\Sigma} = x_d + x_T + x_L/2$; $x_{q\Sigma} = x_q + x_T + x_L/2$; $x'_{d\Sigma} = x'_d + x_T + x_L/2$.

According to the right-inversion theory [7], [8], we have

$$\begin{aligned}\dot{y} &= \frac{\partial V_t}{\partial E'_q}\dot{E}'_q + \frac{\partial V_t}{\partial \delta}\dot{\delta} \\ &= \frac{1}{2y}\left\{\frac{2\left(E'_q\left(x'_{d\Sigma}-x'_d\right)+V_S x'_{d\Sigma}\cos\delta\right)\left(x'_{d\Sigma}-x'_d\right)}{\left(x'_{d\Sigma}\right)^2}\left(\frac{-x_{d\Sigma}E'_q}{x'_{d\Sigma}T_{d0}}+\frac{(x_d-x'_d)V_s\cos\delta}{x'_{d\Sigma}T_{d0}}+\frac{u}{T_{d0}}\right)\right. \\ &\quad \left.+\left(\frac{x_q^2V_S^2\sin(2\delta)}{x_{q\Sigma}^2}+\frac{2\left(E'_q\left(x'_{d\Sigma}-x'_d\right)+V_S x'_d\cos\delta\right)V_S x'_d(-\sin\delta)}{\left(x'_{d\Sigma}\right)^2}\right)(\omega-\omega_0)\right\} \\ &= \dot{y}\left(E'_q,\delta,\omega,u\right)\end{aligned} \tag{3}$$

By simple calculation, we can obtain

$$\frac{\partial \dot{y}}{\partial u} = \frac{\left[E'_q\left(x'_{d\Sigma}-x'_d\right)+V_S x'_{d\Sigma}\cos\delta\right]\left(x'_{d\Sigma}-x'_d\right)}{T_{d0}\left(x'_{d\Sigma}\right)^2 y} \neq 0 \ . \tag{4}$$

It follows from (4) that the excitation system described by (1) and (2) is right invertible, and in theory its right inversion can be described by

$$u = U_{E'_q} = \phi\left(E'_q,\delta,\omega,v\right) \ , \tag{5}$$

where $v = \dot{y} = \dot{V}_t$ is the input of the right inversion. The excitation system can be linearized by cascading the right inversion (5) with the excitation system, at the same time obtaining a one-order linear system s^{-1}, as shown in Fig.2.

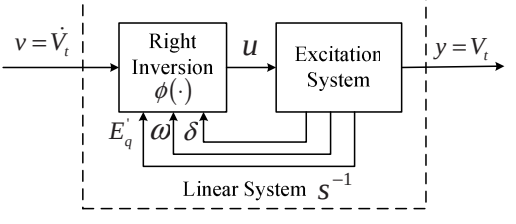


Fig. 2. The linearization of excitation system based on the right inversion

From (5), it is clear to see that the implementation of the right inversion requires the feedback of the full state E'_q, δ, ω. In view that E'_q, δ are directly immeasurable, they can not be directly used to implement the right inversion.

In our previous work [9], we have proposed a soft-sensing method based on the left-inversion, which is capable of estimating the directly immeasurable states on-line.

Here, we will use such a soft-sensing method to estimate the directly immeasurable states E'_q, δ.

According to the soft-sensing theory based on left-inversion [9], we have

$$\dot{\omega} = \frac{\omega_0}{H} P_{m0} - \frac{D}{H}(\omega - \omega_0) - \frac{\omega_0 E'_q V_s}{H x'_{d\Sigma}} \sin\delta - \frac{\omega_0 V_s^2}{2H}\left(\frac{x'_d - x_q}{x'_{d\Sigma} x_{q\Sigma}}\right)\sin 2\delta \ . \tag{6}$$

By calculation, we can obtain the following Jacobian matrix

$$\partial(V_t, \dot{\omega})^T \Big/ \partial\left(\delta, E_q^{'}\right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} , \tag{7}$$

where

$$\begin{aligned}
A_{11} &= \frac{1}{2V_t}\left(\frac{x_q^2 V_S^2 \sin(2\delta)}{x_{q\Sigma}^2} + \frac{2\left(E'_q\left(x'_{d\Sigma} - x_d^{'}\right) + V_S x'_d \cos\delta\right) V_S x'_d \left(-\sin\delta\right)}{\left(x'_{d\Sigma}\right)^2}\right) \\
A_{12} &= \frac{1}{2V_t}\left(\frac{2\left(E'_q\left(x'_{d\Sigma} - x'_d\right) + V_S x'_d \cos\delta\right)\left(x'_{d\Sigma} - x'_d\right)}{\left(x'_{d\Sigma}\right)^2}\right) \\
A_{21} &= -\frac{\omega_0 E'_q V_s}{H x'_{d\Sigma}} \cos\delta - \frac{\omega_0 V_s^2}{H}\left(\frac{x'_d - x_q}{x'_{d\Sigma} x_{q\Sigma}}\right)\cos 2\delta \\
A_{22} &= -\frac{\omega_0 V_s}{H x'_{d\Sigma}} \sin\delta
\end{aligned} \ . \tag{8}$$

In the operating range of excitation system ($0 < \delta < \pi$), we have $rank\left(\partial(V_t, \dot{\omega})^T \Big/ \partial\left(\delta, E_q^{'}\right)\right) = 2$. Therefore, according to the results in [9] and the well-known Inverse Function Theorem, the left-inversion used to estimate δ, E'_q exists and can be expressed by

$$(\delta, E'_q)^T = \hat{\boldsymbol{\phi}}(V_t, \dot{\omega}, \omega) \ . \tag{9}$$

Substituting the left-inversion (9) into the right-inversion (5) forms the following so-called combined-inversion

$$u = \phi\left(\hat{\boldsymbol{\phi}}(V_t, \dot{\omega}, \omega), \omega, v\right) := \bar{\phi}\left(\omega, \dot{\omega}, V_t, v\right) \ . \tag{10}$$

where $v = \dot{y} = \dot{V}_t$ is the input of the combined-inversion.

3 ANN Combined-Inversion

Due to the complexity of the nonlinear system, the explicit expressions of both the right-inversion $\phi\left(E'_q, \delta, \omega, v\right)$ and the left-inversion $\hat{\boldsymbol{\phi}}(V_t, \dot{\omega}, \omega)$ are very difficult to

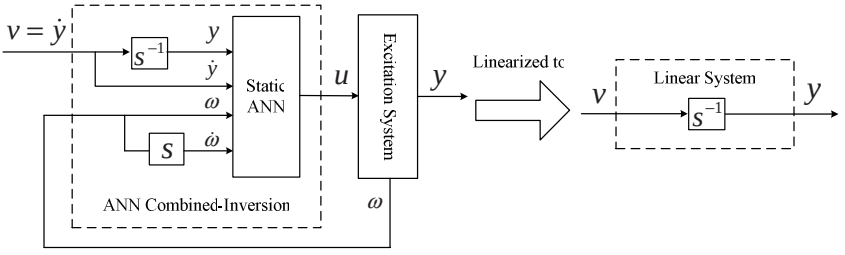


Fig. 3. The linearization of excitation system based on the ANN combined-inversion

obtain although they exist in theory according to the Inverse Function Theorem. This furthermore results in that the explicit expression of the combined-inversion (10) is also difficult to obtain.

To overcome this difficulty, a static ANN is used to approximate the combined-inversion $\bar{\phi}(\cdot)$ due to its strong ability approximating a nonlinear function, thus obtaining the ANN combined-inversion as shown in Fig. 3.

4 Simulation Results

The parameters used in this simulation are as follows: $\omega_0 = 100\pi$, $V_s = 1.0$, $P_{m0} = 0.5$, $x'_d = 0.265$, $x_L = 1.46$, $x_d = 2.156$, $x_q = 2.101$, $x_T = 0.1$, $D = 5$, $H = 8$, $T_{d0} = 10$.

To obtain the ANN accurate enough to approximate $\bar{\phi}(\cdot)$, it is necessary to acquire the training data which should contain enough information of the combined-inversion $\bar{\phi}(\cdot)$. To do so, here the actuating signals are chosen as the square wave signal as shown in Fig. 4.

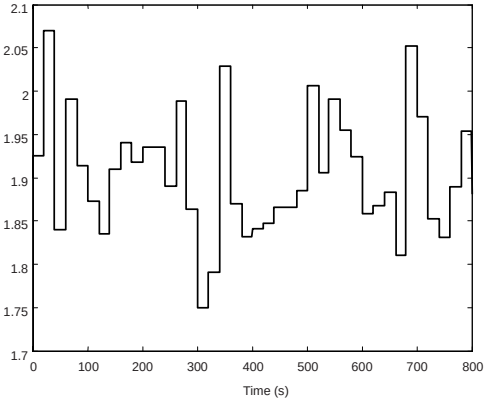


Fig. 4. The actuating signals (square waves)

The data used to train the ANN can be obtained by impacting the above actuating signals to the excitation system. From (10), we can see that the input data of the ANN are $\dot{V}_t, \dot{\omega}, V_t, \omega$ and the output data are u. The data of u, V_t, ω can be obtained directly, and the data of $\dot{V}_t, \dot{\omega}$ can be obtained by using the 5-point numerical derivative algorithm.

The static ANN adopts feed-forward BP network with the structure of 4-10-1 with "tan sigmoid" transfer function on the nodes of the hidden layer, and "linear" transfer function on the node of output layer. The static ANN is trained with the Levenberg-Marquardt training algorithm. After iteratively training the ANN 300 times, the ultimate training Mean Squared Errors (MSEs) is reduced to 9.99933×10^{-10}. The trained ANN-inversion is tested with the data not used for training. The test results show that the generalization of the ANN-inversion is appropriate for actual application.

An ANN right-inversion is also designed according to the expression $u = \phi\left(E_q', \delta, \omega, v\right)$ in order to compare the results between the ANN right-inversion and the ANN combined-inversion. The ANN right-inversion is implemented by straightforward using the states ω, δ, E_q'. Clearly, this can be realized in simulation but not in practice.

Following the Fig.3 and choosing the step signal as the input v, the simulation is performed. Clearly, the excitation system can be linearized and a one-order linear system s^{-1} is obtained.

For the convenience of comparison, Fig.5 shows the linearizing results of ANN right-inversion, those of ANN combined-inversion, and the ideal results which just are the output response of a real one-order integral system.

From the Fig. 5, it is easy to see that the ANN combined-inversion is of excellent performance of linearizing the nonlinear system (1).

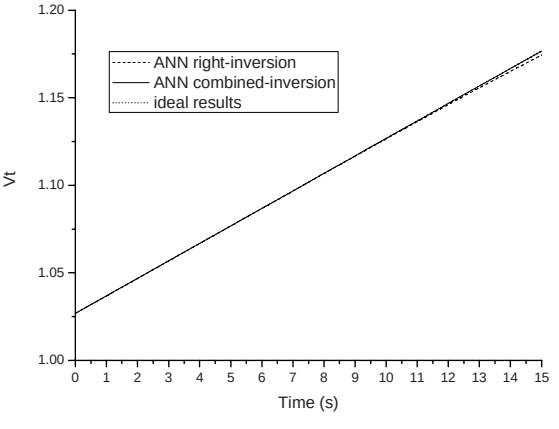


Fig. 5. The linearizing effects of the excitation system

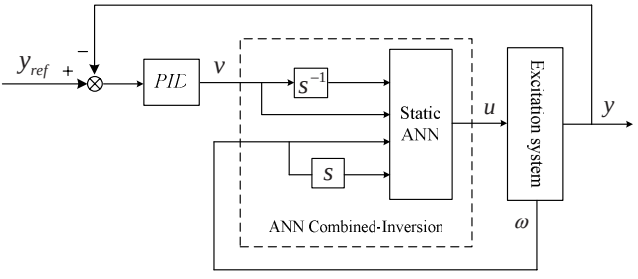


Fig. 6. The ultimate closed control of the excitation system

After being linearized by the ANN combined-inversion, the ultimate closed control of the excitation system can be completed by using the traditional *PID* controller, as shown in Fig. 6 where y_{ref} denotes the reference input of V_t.

To demonstrate the control effectiveness of the ANN combined-inversion in the presence of fault, here consider a typical short-circuit fault of one line (three-phase) in the double transmission lines. At the beginning the system operates with double lines in a stable state; at 1.0s a three-phase symmetrical short-circuit to ground on one of the lines occurs; at 1.15s the faulted line is cut-off and then the system operates at single-line mode; at 1.35 s the faulted line is switched-on again; finally the system remains the double lines transmission.

The simulation results of both the ANN right-inversion and the ANN combined-inversion are shown in Fig.7, from which it can be seen that the ANN combined-inversion has excellent control performance which are almost identical with those of the ANN right-inversion.

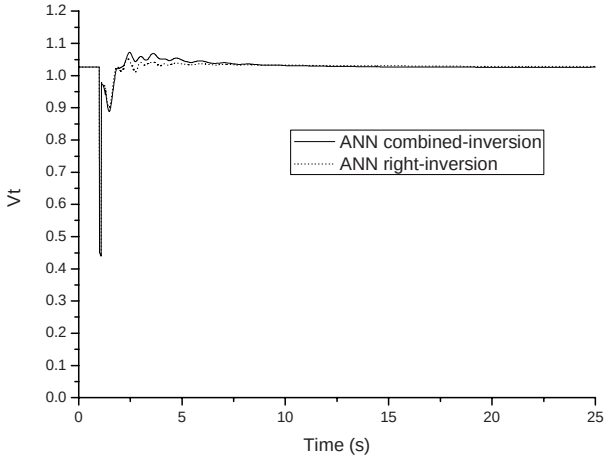


Fig. 7. Control effects of ANN combined-inversion

5 Conclusions

The ANN combined-inversion of the excitation system implemented only by the feedback of measurable variables is capable of not only linearizing the excitation system but also addressing the problem of the feedback of immeasurable states in traditional right-inversion. This makes it more practical to use in engineering. The ANN combined-inversion has excellent control effects.

References

1. Kundur, P.: Power System Stability and Control. McGraw Hill, New York (1994)
2. Machowski, J., Bialek, J.W., Bumby, J.R.: Power System Dynamics and Stability. John Wiley & Sons, New York (1997)
3. Ray, P.S., Duttagupta, P.B., Bhakta, P.: Coordinated Multimachine PSS Design using both Speed and Electric Power. IEE Proc. –Gener. Transm. Distrib. 142(5), 503–510 (1995)
4. Wang, H.F., Swift, F.J., Li, M.: Comparison of Modal Controllability between FACTS-based Stabilisers and PSS in Increasing the Oscillation Stability of Multimachine Power Systems. IEE Proc. –Gener. Transm. Distrib. 143(6), 575–581 (1996)
5. Dai, X., Zhang, K., Zhang, T., Lu, X.: ANN Generalised Inversion Control of Turbo-Generator Governor. IEE Proceedings- Generation, Transmission and Distribution 151(3), 327–333 (2004)
6. Dai, X., He, D., Zhang, T., Zhang, K.: ANN Generalized Inversion for the Linearization and Decoupling Control of Nonlinear Systems. IEE Proceedings- Control Theory and Applications 150(3), 267–277 (2003)
7. Li, C., Feng, Y.: Inversion based Nonlinear Control Method for Multi-variable System. Tsinghua University Press, Beijing (1991)
8. Dai, X., He, D., Zhang, X.: MIMO System Invertibility and Decoupling Control Strategies based on ANN th-order Inversion. IEE Proceedings-Control Theory and Applications 148(2), 125–136 (2001)
9. Dai, X., Wang, W., Ding, Y., Sun, Z.: Assumed Inherent Sensor Inversion Based ANN Dynamic Soft-sensing Method and Its Application in Erythromycin Fermentation Process. Computers and Chemical Engineering 30(8), 1203–1225 (2006)

Learning Action Models with Quantified Conditional Effects for Software Requirement Specification

Hankui Zhuo[1], Lei Li[1], Qiang Yang[2], and Rui Bian[1]

[1] Software Research Institute of Sun Yat-Sen University, China
zhuohank@gmail.com
[2] Hong Kong University of Science and Technology, Hong Kong
qyang@cse.ust.hk

Abstract. Software requirement specification (SRS) is an important step in software engineering. Extracting requirement specification from an application field is a difficult task. In this paper, we consider software requirement as a problem to be solved by intelligent planning. To do this, one of the difficult things is how to represent the domain, since the software requirement has a feature of changeability. Thus, we divide the work into two tasks: the first one is to describe an incomplete domain of software requirement using PDDL(Planning Domains Definition Language) [7]; the second one is to complete the domain by learning from plan samples which is extracted from business processes. We modify the tool of [9] to learn action models with quantified conditional effects, which is the second task. In this way, people only need to do the first task and extract plan samples, which means the efforts of human beings are saved. At the end of the paper, we give our experiment result to show the efficiency of our method.

Keywords: PDDL, SRS, intelligent planning.

1 Introduction

The target of software requirement engineering is to describe an action set[5], which bridges the software system and business process, which is also called SRS. Presently, there are three kinds of SRS [6]: informal specification, semi-formal specification and formal specification. These SRSs, however, do not have the consideration of the lack of solver for the current formal specification. Thus, [6] describes requirement with Answer Set Programming(ASP). It can use ASP solver for requirement validation. It still, however, demands people smart enough to attain SRS with ASP. Inspired by [1], we attempt to describe SRS with PDDL(In this paper, it just means STRIPS with quantified conditional effects)[7], which means we need to write a planning domain. It is time-consuming and toiling. Thus, it makes sense to design an approach to help describe a domain. [1] builds a system called ARMS(*Action-Relation Modelling System*) to learn action model from plan samples with medial states unknown. Inspired by ARMS, a first try

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 874–881, 2008.
© Springer-Verlag Berlin Heidelberg 2008

to apply intelligent planning in software engineering is done by [9]. This application is, however, limited by the feature of ARMS, which is *only STRIPS action models* can be learned.

In software engineering, models with more expressive ability are required. For instance, software engineering has different conditional branches, which means conditional effects are needed in action models; generally, software engineering often refers to many objects in a procedure, that's to say using quantifiers in action models will help to represent SRS concisely. Thus, in this paper, we move on a step to build a tool Learning Action Models from Plan Samples (LAMPS , note that we use the same name as [9]) to learn action models with quantified conditional effects. Learning such a complex action models is more difficult than learning action models in STRIPS description, since building constraints and their corresponding weights are much more difficult. After we build LAMPS, we apply it to software engineering to help extract SRS.

With a set of action schemas and a set of plan traces as input, LAMPS first builds constraints according to action schemas and associates each constraints with a weight by scanning plan traces. After that, LAMPS finds an assignment (true/false) to all the propositional variables to maximally satisfy all the constraints. Finally, LAMPS converts all the propositional variables to action models (as output) according to their assignments. Thus, to apply LAMPS to SRS, it requires software engineers to give action schemas and extract plan traces from real engineering domain. After that, LAMPS will output action models to describe SRS.

By doing this, there are two main advantages we can get: firstly, giving action schemas and plan traces are much easier than describing SRS completely(that is what current formal specification language do), and running the LAMPS is totally automatic, which means the human efforts are saved; secondly, it has planning solvers with PDDL, which makes automatic validation of SRS be possible. In the field of software engineering, automatic validation of SRS is important and very difficult.

The rest of the paper is organized as follows. The next section discusses related work in more detail. After that, the action model learning section describes the algorithm about how to learn action models, which is followed by the extracting SRS section. Finally, the experiment result and make a conclusion are given.

2 Related Work

Learning action model is an important problem in intelligent planning. As a result, many approaches have been built to solve the problem. In [4,3], a partially known action model is improved using knowledge of intermediate states between pairs of actions in a plan.

[1] builds a tool called ARMS to learn action models with medial states unknown. It builds three kinds of constraints from plan samples: predicate constraints, action constraints and plan constraints. After that, it associates each constraint with a weight by scanning all the plan traces. In the end, it solves

these constraints with the solver of weighted-MAXSAT[8] to maximally satisfy all the constraints and outputs action models. Based on ARMS, [9] takes a first try to apply intelligent planning in software engineering. It first requires software engineers to give action schemas and plan traces extracted from real engineering domain. Then, it explores LAMPS (a revised version of ARMS) to output action models to describe SRS.

The idea using LAMPS in SRS is inspired by [6]. In the traditional formal specifications, there isn't powerful solver to interpret the result. For instance, when we build a Petri Network to describe SRS, there isn't general solver to interpret the Petri Network. In contrast, [6] converts SRS to formulas, and then solves them by ASP solver. It makes the formal requirement validation possible using ASP solver. Similarly, in this paper we represent SRS to action models with PDDL. Compared to [6], we can use action model learning tool LAMPS to reduce our effort to describe SRS.

3 Learning Action Models

A plan is an action sequence $(a_0, a_1, \ldots, a_n)$ which makes a projection from the initial state s_0 to the goal state g. A set of plan samples is defined as $\mathcal{T} = \{T|T = (s_0, a_0, a_1, \ldots, a_n, g)\}$. A set of action schemas is defined as $\mathcal{A} = \{\bar{a}|\bar{a}$ is composed of action name and parameters$\}$.

Our learning problem is: given a set of plan samples $\mathcal{T}$ and a set of action schemas $\mathcal{A}$, the algorithm outputs preconditions and effects of each action schema in $\mathcal{A}$. We call as a conditional effect $when\langle conditons\rangle\langle effects\rangle$, where $\langle conditions\rangle$ are the conditions of a conditional effect, and $\langle effects\rangle$ are the effects of a conditional effect. An example of our learning problem is shown in Tables 1 which includes action schemas, plan samples from *Briefcase* and action models as output. To learn the action model, we modify the approach of [9] as follows.

==========

Step 1: An incomplete domain $\bar{\Sigma}$ and a set of plan samples is given as input.
Step 2: For every action a, find the Possible Predicate Set(PPS) related to a, and denote this set as P_a.
Step 3: For every action a, build action constraints with P_a.
Step 4: Build plan constraints and predicate constraints using plan samples.
Step 5: Build constraints for conditional effects.
Step 6: Build model constraints based on the learned action models.
Step 7: Let these constraints as the input of Weighted MAXSAT, and convert the output of Weighted MAXSAT to temporary action models [1].
Step 8: Find the most credible action as a learned action model and update plan samples.
Step 9: If all the actions are learned, then stop, else go to step 4.

==========

In the algorithm, the detail of Step 2 to 4 and Step 6 to 8 can be seen from [9]. Thus, we just give the detail of Step 5 here. We find out predicate-action

Table 1. Domain description of *briefcase*; Input: plan traces from domain *briefcase*; Output: the learned action models of the domain *briefcase*

Domain description: *briefcase*	
predicates:	(at ?y - portable ?x - location) (in ?x - portable) (is-at ?x - location)
action:	move(?m - location ?l - location) take-out (?x - portable) put-in (?x - portable ?l - location)

Inputs			**Output**	
	plan trace 1	plan trace 2	action:	move(?m - location ?l - location)
Initial	(is-at l1) (at o1 l1) (at o2 l2)	(is-at home) (at o1 l1)	pre:	(is-at ?m)
Actions	(put-in o1 l1) (move l1 l2) (put-in o2 l2) (move l2 home)	(move home l1) (put-in o1 l1) (move l1 home)	effect:	(and(is-at ?l)(not (is-at ?m)) (forall (?x - portable) (when (in ?x) ((and(at ?x ?l)(not(at ?x ?m)))))))
Goal	(is-at home) (at o1 home) (at o2 home)	(is-at home) (at o1 home)	... ...	

pair from last actions and goals with frequent set mining. For a predicate-action pair $< a, p >$, there is at least one parameter in p, which is not in a. We guess an action's effect according to predicate-action pair's frequency. We make a assumption: For all predicates in the conditions of a conditional effect, their parameters are either in the action or in the effects of the conditional effect and at least one parameter in the effects of the conditional effect. Based on the assumption, we build a Possible-Predicate Set for conditions of a conditional effect (we denote this set as CPPS). We build constraints based on CPPS as follows: $ADD(c) \cap PRE(c) = \emptyset$, $ADD(c) \cap DEL(c) = \emptyset$, $PRE(c) \neq \emptyset$ and $ADD(c) \cup DEL(c) \neq \emptyset$, where c is a conditional effect, $ADD(c)$ is the adding effects of c, $DEL(c)$ is the deleting effects of c, and $PRE(c)$ is the conditions of c. Furthermore, for every predicate p in CPPS, we guess p is a condition of a conditional effect. Then, we build the constraint: $\forall p.(p \in CPPS \Rightarrow p \in PRE(c))$. In a plan sample t, if p is the only predicate which is (1) in goal but not in t's initial state; (2) generated by some when-clause; (3) that a is the last action that has a when-clause, denoted as c, to generate p, then we guess p is generated by a's conditional effect. That's to say, the conditions of the conditional effect are satisfied. Furthermore, if a is also the first action in t, we have the following constraint: $\forall p \in CPPS(c) \wedge p \in s_0 \Rightarrow p \in PRE(c)$, where s_0 is a initial state, $CPPS(c)$ is the possible-predicates of c. If p is the last predicate in goal which is not included by initial state, and the first action of the plan, which is not the last action in the plan, has a conditional effect c to generate p, then there is at least one predicate in c's condition but not in the initial state. If it is not

the case, the actions except the first action are not necessary. Then we build the constraint: $\exists p \in CPPS(c) \wedge p \in PRE(c) \Rightarrow p \notin s_0$. After we get these constraints for conditional effects, we associate each constraint with a weight, which is its occurrence in the plan samples.

4 Extracting Software Requirement Specification

To explore LAMPS in SRS, the first thing we have to do is to transform the software requirement into an incomplete domain by PDDL. Here, we can assume an incomplete domain is composed of three parts: type definition, predicate definition and incomplete action model definition. Therefore, we have to attain this three parts. To do this, we mainly base on three basic ideas.

4.1 Definition of Types and Predicates

Our first basic idea is: *a whole story can be stated as who do something to what, when and where.*

Example 1. *the sentence "I read book in the lab yesterday", the 'I', 'read', 'book', 'in the lab' and 'yesterday' mean 'who', 'do something', 'what', 'where' and 'when' respectively.*
Similarly, we extract objects from software requirement as follows: subject, object, environment and time, which are who, what, where and when respectively. For simplicity, we don't discuss time object here. Therefore, we can have the following type definition:
(:type
subject object environment – objects
subject1 subject2 ... subjectn – subject
object1 object2 ... objectn – object
environment1 environment2 ... environmentn – environment
)
We can comprehend the definition by grasping PDDL's type definition. Of course, it's possible that every subtype can have its own subtypes. For example, subject1 is subject's subtype, and subject11 can be subject1's subtype according to specific application.

Our second basic idea is: what a predicate represents is some object's feature or some relation between two objects. It is possible that there are predicates represent some relations among more than two objects, but we don't concern that in this paper, which is left for our further work. The detail can be seen in Example 2.

4.2 Incomplete Action Model Definition

An incomplete action model is composed of action name and parameters. Hence we get an incomplete action model by getting action name and parameters respectively. Firstly, in a specific SRS there are some business operator. We get an

action name according to one business operator. And then, we make an assumption that every object the business operator concerns can be action's parameters. These objects can be: (1) business operator's subjects, (2) business operator's objects, (3) business operator's resultant object, (4) the environment object in which business operator happens. Also, we'll see Example 2 to make it clearer.

Example 2. *Bookstore ordering system can be separated by three levels:*
1. client ordering: a client orders his books from base store based on list of book, and produces an initial ordering form.
2. base store ordering: a staff of base store orders his books from province store based on initial ordering forms, and produces base ordering forms.
3.province store ordering: a staff of province store orders his book from supplier, and produces province ordering forms.
The objects that the system concerns is as shown in Table 2. After that, types can be defined accordingly, which are shown in the same table. Next, according to Sect.4.2, we find out two types of predicates respectively. The result is also shown in Table 2. Finally, we build an incomplete action model. There are three business processes: client ordering, base store ordering and province store ordering. They are also three business operators. Therefore, we give three action names ClientOrder*,* BaseOrder *and* ProvinceOrder *correspondingly. And we can attain parameters according to Sect.4.3. The resultant incomplete action model is: (ClientOrder ?x - staff ?y - bstore), (BaseOrder ?x - staff ?y - pstore), (ProvinceOrder ?x - staff ?y - sdep). Since executing these actions is a process, we need an action to stop the process. We call this action as* BookDone *and its related objects are the above action's objects except the environment objects. That is (OrderDone ?x - staff).*

5 Experiment

For simplicity, we do the experiment with the bookstore ordering system of Example 2. Our purpose is to extract the system's requirement specification described by PDDL. We extract 40 plan samples from the real application. With these samples and the incomplete domain as input, LAMPS outputs action models as shown in Table 3. According to business experience, the result of Table 3 is adapted. The part which is italicized and emphasized means it must be deleted. The part which is emphasized means it must be added. The resultant meanings of the action models are: the process of client ordering (*ClientOrder*) can be executed, only when the staff ?x who belongs to a base store is free, and initial ordering form ?a is in its initial state; as a result of the execution, the staff ?x is busy, and forall *book list* ?b is being dealt with if it is in waiting state. Similarly, the process of the *BaseOrder* and *ProvinceOrder* have the same meaning as *ClientOrder*. The *OrderDone* action means: when the staff x is busy in dealing with the form ?a and ?b, then the action can be executed and produce the results: staff ?x is free, the form ?a is dealt with and ?b is waiting for being dealt with.

Table 2. Definitions of bookstore ordering system

Objects in bookstore ordering system		
subject	*object*	*environment*
staff	list of book (lbook)	base store (bstore)
	initial ordering form (ioform)	province store (pstore)
	base ordering form (boform)	supplier'department (sdep)
	province ordering form (poform)	

Type definition of bookstore ordering system	
types:	staff form department - object
	lbook ioform boform poform - form
	bstore pstore sdep - department

Predicates and their meanings	
predicates	meanings
(init ?x - form)	the form ?x is in initial state
(waiting ?x - form)	the form ?x is in waiting state
(dealt ?x - form)	the form ?x is dealt
(free ?x - staff)	the staff ?x is free
(busy ?x - staff)	the staff ?x is busy
(doing ?x - staff ?y - form)	the staff ?x is dealing with the form ?y
(belong ?x - staff ?y - department)	the staff ?x belongs to department ?y

Incomplete action models	
actions:	(ClientOrder ?x - staff ?y - bstore ?a - lbook ?b - ioform)
	(BaseOrder ?x - staff ?y - pstore ?a - ioform ?b - boform)
	(ProvinceOrder ?x - staff ?y - sdep ?a - boform ?b - poform)
	(OrderDone ?x - staff ?a - form ?b - form)

Table 3. Action Models of Bookstore Ordering System

(ClientOrder ?x - staff ?y - bstore ?a - ioform)	
precondition:	(free ?x) (init ?a) (belong ?x ?y)
effect:	(:and (busy ?x) (not (free ?x)) (not (init ?a)) ***(doing ?x ?a)*** (forall ?b - lbook (when (waiting ?b) (and (doing ?x ?b)(not (waiting ?b))))))
(BaseOrder ?x - staff ?y - pstore ?a - boform)	
precondition:	free(?x) init(?a) belong(?x,?y)
effect:	(:and (busy ?x) (not (free ?x)) (not (init ?a)) (doing ?x ?a) (forall ?b - ioform (when (waiting ?b) (and (doing ?x ?b)**(not (waiting ?b))**)))))
(ProvinceOrder ?x - staff ?y - sdep ?a - poform)	
precondition:	free(?x) init(?a) belong(?x,?y)
effect:	(:and (busy ?x) ***(not (free ?x))*** (not (init ?a)) (doing ?x ?a) (forall ?b - ioform (when (waiting ?b) (and (doing ?x ?b)**(not (waiting ?b))**)))))
(OrderDone ?x - staff ?a - form ?b - form)	
precondition:	***(dealt ?b) (waiting ?a)*** (busy ?x) (doing ?x ?a) (doing ?x ?b)
effect:	(dealt ?a) (waiting ?b) (free ?x) ***(not (dealt ?b))*** (not (busy ?x)) **(not (doing ?x ?a)) (not (doing ?x ?b))**

From the experiment, we can see that: (1) LAMPS helps to extract action models, or software requirement, but manual adjustment is needed. (2) It's hard to discuss how good or bad is the experiment result, which depends on the specific application. (3) It is somehow easier to adjust an existing action model than to build an action model. That is to say, it's easier to extract SRS with LAMPS than to extract SRS all by manual effort. But it's hard to say how much effort it saves.

6 Conclusion

From this paper, we make see that, firstly, LAMPS can learn action models with conditional effect and universal quantifiers, which moves a step in learning action models, compared to ARMS. Secondly, we take a first try by learning to apply intelligent planning to software requirement specification. We do not discuss how much effort it saves to attain SRS quantitatively, but we are sure about that our method is semiautomated, and it makes sense to describe SRS with planning domain, for example, for automated software requirement validation. Finally, we also know that LAMPS cannot learn action models with respect to resource planning, which is very important in SRS. Thus, extending LAMPS to learning resource will be our next work.

References

1. Yang, Q., Wu, K.H., Jiang, Y.F.: Learning Action Models from Plan Examples with Incomplete Knowledge. In: Proceedings of the 2005 International Conference on Automated Planning and Scheduling (ICAPS 2005), Monterey, CA, USA, June, pp. 241–250 (2005)
2. Blythe, J., Kim, J., Ramachandran, S., Gil, Y.: An Integrated Environment for Knowledge Acquisition. In: Proceedings of the 2001 International Conference on Intelligent User Interfaces (IUI 2001), Santa Fe, NM, pp. 13–20 (2001)
3. Gil, Y.: Learning by Experimentation: Incremental Refinement of Incomplete Planning Domains. In: Eleventh Intl. Conf. on Machine Learning, pp. 87–95 (1994)
4. Oates, T., Cohen, P.R.: Searching for Planning Operators with Context-dependent and Probabilistic Effects. In: Proceedings of the Thirteenth National Conference on AI (AAAI 1996), Portland, OR, pp. 865–868 (1996)
5. Sommerville, I.: Software Engineering. Addison-Wesley (2000)
6. Wan, H., Zheng, Y.X., Li, L.: Software Requirement Specification Based on Answer Sets Semantics and Subject-Predicate-Object. In: The 8th International Conference for Young Computer Scientists (ICYCS 2005), Beijing, China, September 20-22 (2005)
7. Fox, M., Long, D.: PDDL2.1: an Extension to PDDL for Expressing Temporal Planning Domains. Journal of Artificial Intelligence Research 20, 61–124 (2003)
8. Borchers, B., Furman, J.: A Two-phase Exact Algorithm for MAX-SAT and Weighted MAX-SAT Problems. Journal of Combinatorial Optimization 2(4), 299–306 (1999)
9. Zhuo, H.K., Li, L., Bian, R., Wan, H.: Requirement Specification Based on Action Model Learning. In: Huang, D.-S., Heutte, L., Loog, M. (eds.) ICIC 2007. LNCS, vol. 4681, pp. 565–574. Springer, Heidelberg (2007)

A New Method to Determine Evidence Discounting Coefficient

Wen Jiang*, An Zhang, and Qi Yang

School of Electronics and Information, Northwestern Polytechnical University,
Xi'an, 710072, P. R. China
jiangwenpaper@yahoo.cn

Abstract. Data fusion technology is widely used in automatic target recognition systems to improve the efficiency. In the framework of evidence, information fusion relies on the use of a combination rule allowing the belief function to be combined. However, Dempster's rule of combination is a poor solution for the management of the conflict between the various information sources. It is proved that, if the false evidence can be correctly selected, Dempster's rule can deal with highly conflicting evidence combination efficiently. However, how to determine the false evidence is an open issue. In this paper, a novel method to determine the discounting coefficient is proposed. First, the distance function between bodies of evidence is introduced to express the degree of conflict degree. Then, the confidence lever of each piece of evidence is obtained to reflect the reliability of each information source to some degree. The discounting coefficient can be determined finally through the relative credibility of sensor reports. The numerical example of multi-sensor fusion target recognition based on DS theory is shown to illustrate the efficiency of the presented approach.

1 Introduction

Information fusion is widely used in many application fields, such as image processing and analysis, classification or target tracking. In the framework of evidence theory, information fusion relies on the use of combination rule allowing the belief functions for the different propositions to be combined. A crucial role in evidence theory is played by Dempster's combination rule that has several interesting mathematical properties such as commutativity and associativity [1][2].

However, combining belief functions with this operator implies normalizing the results by scaling them proportionally to the conflicting mass in order to keep some basic properties. In [3], Zadeh has underlined that this normalization involves counter-intuitive behavious. In order to solve the problem of conflict management, Yager [4], Smets[5], Dubois[6]and more recently Lefevre[7] have proposed other combination rules. However, these rules have more or less satisfactory behavious.

One of the most important reasons for evidence conflicting is that the aberrant measurement given by a sensor. In fact, an abnormal measurement can generate a conflicting mass during the combination step [7]. Hence, in the procedure of evidence combination,

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 882–887, 2008.
© Springer-Verlag Berlin Heidelberg 2008

it is desired to consider the reliability of each sources of information. In [8], the problem of false evidence processing in information fusion is discussed. It is shown that the combination will be better if we can efficiently deal with false evidence. Discounting method is proved to be a good alternative to deal with evidence combination in highly conflicting environments[7]. For example, the discounting coefficient of a piece of evidence is low if its reliability is less than other evidence in data fusion systems. It is obvious that the key issue is how to determine the discounting coefficient of each evidence. In this paper, a new method to determine discounting coefficient is proposed.

2 Dempster Shafer Evidence Theory and Its Problem

Evidence theory first supposes the definition of a set of hypotheses Θ called the *frame of discernmen*t, defined as follows:

$$\Theta = \{H_1, H_2, \ldots, H_N\} \tag{1}$$

It is composed of Nexhaustive and exclusive hypotheses. Form the frame of discernment Θ, let us denote $P(\Theta)$, the *power set* composed with the 2^N propostions $A$of Θ:

$$P(\Theta) = \{\emptyset, \{H_1\}, \{H_2\}, \ldots, \{H_N\}, \{H_1 \cup H_2\}, \{H_1 \cup H_3\} \ldots, \Theta\} \tag{2}$$

where Ø denotes the empty set. The Nsubsets containing only one element are called *singleton*s. A *BPA* is a function from $P(\Theta)$ to [0,1] defined by :

$$m : P(\Theta) \rightarrow [0, 1]$$

$$A \mapsto m(A) \tag{3}$$

and which satisfies the following conditions:

$$\sum_{A \in P(\Theta)} m(A) = 1 \quad m(\emptyset) = 0 \tag{4}$$

In the case of imperfect data (uncertain, imprecise and incomplete), fusion is an interesting solution to obtain more relevant information. Evidence theory offers appropriate aggregation tools. From the basic belief assignment denoted m_i obtained for each information source S_i, it is possible to use a combination rule in order to provide combined masses synthesizing the knowledge of the different sources. Dempsters rule of combination (also called orthogonal sum), noted by $m = m_1 \oplus m_2$, is the first one within the framework of evidence theory which can combine two *BPAs* m_1 and m_2 to yield a new *BPA* :

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - k} \tag{5}$$

with

$$k = \sum\nolimits_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{6}$$

where $k$is a normalization constant, called *conflict* because it measures the degree f conflict between m_1 and m_2 $k = 0$ corresponds to the absence of conflict between m_1 and m_2, whereas $k = 1$ implies complete contradiction between m_1 and m_2. The belief function resulting from the combination of J information sources S_J defined as

$$m = m_1 \oplus m_2 \cdots \oplus m_j \cdots\cdots \oplus m_J \tag{7}$$

When conflicting evidence is present, Dempster's rule for combining beliefs often produces counter-intuitive results. A typical example given by Zadeh [3]is as follows:

Example 1. Consider a situation in which we have two belief structures m_1 and m_2 as follows:

$$m_1(a) = 0.99 \quad m_1(b) = 0.01$$

$$m_2(b) = 0.01 \quad m_2(c) = 0.99$$

Application of the Dempster rule yields

$$m(b) = \frac{0.0001}{1-k} = \frac{0.0001}{1-0.9999} = 1$$

Thus it can be seen that while m_1 and m_2 afford little support to b, the results affords complete support to b. This appears somewhat counterintuitive. For the illogical aspects mentioned above, conflict management in belief functions is a very important problem and has been already studied in the past.

To solve these problems, several alternatives to the normalization process have been proposed [4-8]. One of the efficient method is the use discounting coefficient. Here we introduce briefly.

In [8],the problem of false evidence processing in information fusion is discussed. It is shown that the combination will be better if we can efficiently deal with false evidence. For example, we can combine the negative evidence of the false evidence to decrease the influence of false evidence. It is obvious that the key step of the false evidence processing is to select the false evidence correctly. The following section will detail the algorithm to correctly select the false evidence.

$$\begin{cases} m^{\alpha}(\Theta) = \alpha m(\Theta) + 1 - \alpha, \\ m^{\alpha}(A) = \alpha m(A), \quad \forall A \subset \Theta \end{cases} \tag{8}$$

3 False Evidence Selection Based on Distance

In this section, based on the definition 1-4 of distance measure between bodies of evidence presented by Jousselme [9], a credibility degree coefficient is introduced to show the reliability of each piece of evidence. By setting the threshold, we can select the false evidence whose credibility degree coefficient, or in other words, whose reliability is lower than the value of the threshold.

Unlike other's viewpoint, Jousselme [9]etc give a geometrical interpretation for *BPAs*. The main idea is detailed as follows: Let us call $\mathrm{E}_{P(\Theta)}$, the space generated by all subsets

of Θ. $\mathrm{E}_{P(\Theta)}$ is the a vector space if any linear combination of the object of $\mathrm{E}_{P(\Theta)}$ is in $\mathrm{E}_{P(\Theta)}$. That means

$$\overrightarrow{V} = \sum_{i=1}^{2^N} \alpha_i(A_i) \in \mathrm{E}_{P(\Theta)} \tag{9}$$

where A_i is an element of $P(\Theta)$ and $\alpha_i \in \Re$, being the real set. Then, $A_1, A_2, \ldots, A_{2^N}$ forms a base for $\mathrm{E}_{P(\Theta)}$. Because condition (9) is easily satisfied, we can state that $\mathrm{E}_{P(\Theta)}$ is a vector space. Thus, this will lead to the definition of a *BPA* as a special case of vectors of $\mathrm{E}_{P(\Theta)}$.

Definition 1. Let Θ be a frame of discernment containing Nmutually exclusive and exhaustive hypotheses, and let $\mathrm{E}_{P(\Theta)}$ be the space generated by all subsets of Θ. A basic probability assignments (BPA) is a vector $m$of $\mathrm{E}_{P(\Theta)}$ with coordinates $m(A_i)$ such that

$$\sum_{i=1}^{2^N} m(A_i) = 1 \; and \; m(A_i) \geq 0 \; , i = 1, \ldots, 2^N \; , \; A_i \in P(\Theta) \tag{10}$$

Definition 2. Let m_1 and m_2 be two BPAs on the same frame of discernment Θ, containing Nmutually exclusive and exhaustive hypotheses. The distance between 1 $m$and 2 m, as presented in [9], is:

$$d_{BPA}(m_1, m_2) = \sqrt{\frac{1}{2}\left(\overrightarrow{m_1} - \overrightarrow{m_2}\right)^T \underline{\underline{D}} \left(\overrightarrow{m_1} - \overrightarrow{m_2}\right)} \tag{11}$$

where $\underline{\underline{D}}$ is an $2^N \times 2^N$ matrix whose elements are

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad A, B \in P(\Theta) \tag{12}$$

In general, each piece of evidence has its own reliability. To the decision-maker, he or she often thinks one piece of evidence is more reliable than the other piece of evidence. How to determine the reliability of each evidence is very important, especially when the evidence collected by multi-sensor has high degree of conflict. A reasonable idea to handle this problem is that: If a piece of evidence is supported by other evidence in the system, then the evidence should be more reliable than that of evidence which has high conflicting degree with other evidence. Based on this rule, we given some definitions to describe the credibility degree of evidence.

Definition 3. Assume n pieces of evidence are collected in the system. The number of the there are The similarity measure Sim between m_i and m_j is given as:

$$Sim(m_i, m_j) = 1 - d_{BPA}(m_i, m_j) \quad i, j = 1, 2, \ldots, n \tag{13}$$

Definition 4. The support degree Sup of each piece of evidence is defined as:

$$Sup(m_i) = \sum_{\substack{j=1 \\ j \neq i}}^{2^N} Sim(m_i, m_j) \quad i, j = 1, 2, \ldots, n \tag{14}$$

Definition 5. The credibility degree Crd of each piece of evidence, or the discounting coefficient α_i is defined as:

$$\alpha_i = Crd_i = \frac{Sup(m_i)}{n-1} \quad i = 1, 2, \ldots, n \tag{15}$$

As can be seen in Eq. (15), for a piece of evidence, the less the distance between others, the more the support degree. If a piece of evidence is supported by other evidence greatly, its credibility degree is high and this evidence has more effect on the final combination results. On the contrary, if a piece of evidence is always conflict with other evidence with high degree, its credibility degree is low and this evidence should have less effect on the final combination results.

4 Numerical Example

The following example is illustrated the procedure of the proposed method to select the false evidence.

Example 2 In a multisensor based automatic target recognition system, suppose the real target is A and the system has collected 4 pieces of evidence m_1, m_2, m_3, m_4 as follows:

$$m_1 : m_1(A) = 0.5 \quad m_1(B) = 0.2 \quad m_1(C) = 0.3$$

$$m_2 : m_2(A) = 0 \quad m_2(B) = 0.9 \quad m_2(C) = 0.1$$

$$m_3 : m_3(A) = 0.55 \quad m_3(B) = 0.1 \quad m_3(C) = 0.35$$

$$m_4 : m_4(A) = 0.55 \quad m_4(B) = 0.1 \quad m_4(C) = 0.35$$

According to the definition 4, the credibility degree coefficient of each piece of evidence can be calculated as follows:

$$Sup(m_1) = 1.8491 \quad Sup(m_2) = 0.2299 \quad Sup(m_3) = 2.0195 \quad Sup(m_4) = 2.0195$$

The credibility degree Crd of each piece of evidence can be obtained as follows:

$$Crd(m_1) = 0.6164 \quad Crd(m_2) = 0.0766 \quad Crd(m_3) = 0.6732 \quad Crd(m_4) = 0.6732$$

Hence, according to the discounting coefficient above, the final result can be shown as follows:

$$m(A) = 0.625; m(B) = 0.082; m(C) = 0.217; m(A, B, C) = 0.076$$

Thus, it can be easily to decide that the target is A, though the second evidence highly conflicting with each other. On the other hand, if we use classical Dempster Shafer combination rule, the mass of A is zero, which is not the correct. The simple example used here can easily show the efficiency of the proposed discounting coefficient.

5 Conclusion

One of the main reasons for the poor conflicting evidence management of Dempster's combination operator is that the reliability of each information source is not taken into account. Using discounting coefficient may be useful to handle conflicting evidence combination. However, how to determine the discounting coefficient is not well addressed. In this paper, the distance function of evidence is introduced to measure the support degree of each piece of evidence. Then, the discounting coefficient is determined taking consideration of the credibility of each sensor report. A numerical example in target recognition is used to show the efficiency of the proposed method.

Acknowledgments. Many thanks for Dr. Deng Yong for the constructive communication about the distance on evidence.

References

1. Dempster, A.P.: Upper and Lower Probabilities Induced by a Multi-valued Mapping. J. Annual Math Statist. 38(4), 325–339 (1967)
2. Shafer, G.: A Mathematical Theory of Evidence. Princeton University, Princeton (1976)
3. Zadeh, L.: A Simple View of the Dempster-Shafer Theory of Evidence and Its Implication for the Rule of Combination. J. AI Magazine 7(1), 85–90 (1986)
4. Yager, R.: On the Dempster-Shafer Framework and New Combination Rules. J. Information Science 41(2), 93–137 (1989)
5. Dubois, D., Prade, H.: Representation and Combination of Uncertainty with Belief Functions and Possibility Measures. J. Computational Intelligence (4), 244–264 (1998)
6. Smets, P.: The Combination of Evidence in the Transfer Belief Model. J. IEEE Transaction on Pattern and Machine Intelligence 6(4), 447–458 (1990)
7. Lefevre, E., Colot, O., Vannoorenberghe, P.: Belief Function Combination and Conflict Management. J. Information Fusion 3(3), 149–162 (2002)
8. Tao, Z.X., Liu, Z.Q., Zhang, W.D.: False Evidence Processing in Information Fusion. J. Systems Engineering and Electronics 24(9), 7–9 (2002) (in chinese)
9. Jousselme, A.L., Grenier, D., Bosse, E.: A New Distance between Two Bodies of Evidence. J. Information fusion 2(1), 91–101 (2001)

Feedback Information Expression and Fusion Method for Human-Robot Interaction

Qijie Zhao, Dawei Tu, Jianxia Lu, and Zhihua Huang

School of Mechatronical Engineering and Automation, Shanghai University, Shanghai 200072
{zqj,tdwshu,ljx}@staff.shu.edu.cn,
hzhshu@shu.edu.cn

Abstract. On the basis of analyzing the information fusion method for human-robot interaction, a kind of information feedback structure for human--robot interaction platform has been provided, the source feedback information is handled with multi-layer processed structure, and the processed multi-mode information is integrated with certain interactive rules and knowledge in interaction knowledge database. A set of feedback information expression and fusion method has been presented in information integration process, the abstract information from feedback multimodality is expressed by the situation of task execution process and the stability of feedback modality, and then these information is fused in light of the result of task execution, at last, the fusion result is delivered to user by the most stable feedback modality. Some experiments have been done with the provided methods in the human–robot interaction system, and parts of the experiment results show that user's cognition loads can be decreased by this semantic information fusion and feedback method, moreover, high work efficiency also can be gotten in human-robot interaction.

Keywords: human-robot interaction; feedback information; multi modality; information fusion.

1 Introduction

In the human-robot interaction, the information received from a variety of sensing channels is combined into specific commands or meanings [1], in order to realize the duty request and response between human and machine. Presently, in human-robot interaction field many research works about the information fusion of human-robot interaction have been done [2,3], but these works mostly focused on the information fusion of multiple input channels, and relatively few were about the multiple channel feedback information fusion of the task execution result of machine(robot). The redundant and supplementary feedback information may let the user intuitively and correctly perceive the task execution result. Especially in nurse service robot system, the users are usually the elderly and disabled persons who move with difficulty or have certain barriers in their bodies. Through analyzing the user behavior characteristics, the users will gain better services using the suitable channel and intuitively feedback information. In this paper, starting from the characteristic of the human-robot

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 888–895, 2008.
© Springer-Verlag Berlin Heidelberg 2008

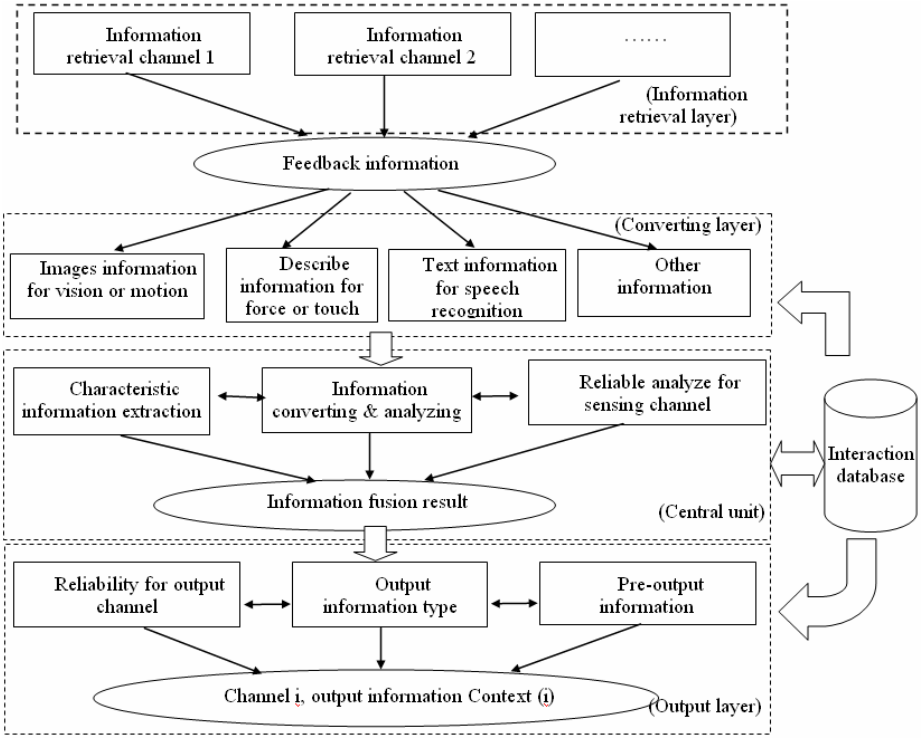


Fig. 1. Information feedback structure

interaction process, the abstract information which is fed back from multiple channels is fused on the semantic layer, and fed back to the user through an intuitively and visual channel, in this way, the user's cognition intensity is reduced, and the human-robot interactive efficiency is enhanced.

2 Information Feedback System

The "human- robot interaction platform" mode has been presented in the human-robot system, the interactive feedback information needs to be integrated through certain fusion rules, and the final semantic information will be intuitively fed back to user through suitable channel of the interactive platform. Information processing procedure is shown in Figure 1, in which the multiple information acquisition channels get the situation of the executing task, moreover, the original status information for both robot and object can be gotten.

The feedback raw information contain types of information, such as image, word and so on, the raw information is transformed into information which can be fused in the transformation intermediate level, and the transformed feedback information will be confirmed in the central processing unit in light of some factors, such as the human-robot interaction knowledge database, the characteristic information extracted

from the raw information, the feedback channel characteristic, and so on. The output intermediate level classify the feedback integrated information according to the characteristics of the output channel and the output information, in order to most intuitively feed back to the terminal user with the most reliability channel.

3 Expressions and Integration Algorithm of the Feedback Information

3.1 Expressions of the Feedback Information

In figure 1, information transformation intermediate level classifies original information fed back from different sensing channels. Where, InputSM[i] express some sensing channel, and SM is the set of multiple sensing channels which are composed of InputSM[i], and then SM can be expressed as:

SM = {InputSM[i]|i=1,2,3,…n} = {vision, moving sensation, force sensation, tactual sensation, speech, network,…}

We can use InputSM[i]_OContext[j] to denote the j-th original content fed back from sensing channel InputSM[i], and use SM_OC to denote the original information types set fed back from all sensing channel, then SM_OC can be expressed as:

SM_OC = {InputSM[i]_OContext[j]|i=1,2,3,…m, j =1,2,…, n} = {picture information | Description information | Text information | Numerical information | …}

The task execution feedback information of robot generally is the status information such as the information about the task executed, the object captured, the goal moved, and so on. Each kind of sensing information is transformed into status information according to knowledge and rules of the knowledge database. We can define the original information InputSM[i]_OContext[j] of the sensing channel InputSM[i] as InputSM[i]_CContext[j], and SM_CC as the information set of the transformed feedback information from multiple channels, then SM_CC can be expressed as:

SM_CC = {InputSM[i]_CContext[j]|i=1,2,…, m, j=1,2,…, n}
= {task finished | task not finished | object grasped | object not grasped | …}

In order to realize the automatic reasoning and comparing of the computing element, the well transformed information set is further processed into numerical or Boolean information which the computing element can compare, such as ”the task finished “or “the task not finished” can be respectively expressed as: Task_Finished = TRUE, Task_Finished = FALSE. Let InputSM[i]_SValue[j] be status value of feedback information corresponding to the InputSM[i]_CContext[j] in transformed information set SM_CC, so the feedback information set is a set only including status information. We can denote SM_SS as this status information set, then:

SM_SS={InputSM[i]_SValue[j]|i=1,2,…,m.j=1,2,..,n}={TRUE| FALSE,… }

In order to distinguishing the efficiency and reliability of different sensing channel in feeding back information, according to the special situation of each sensing channel under different environment, we classify the reliability of each channel to different level. We can define InputSM[i]_StableLevel as the reliability level of the sensing channel InputSM[i], and the reliability grades can be divided into 1,2,3, and so on.

SM_SL is defined as the reliability grade set of the sensing channels, and then SM_SL can be expressed as:

$$SM_SL = \{InputSM[i]_StableLevel | i=1,2,3,\ldots,m\} = \{1, 2, 3, \ldots\}.$$

Stipulating that the bigger value of InputSM[i]_StableLevel, the higher reliability level of the sensing channel .

Let TFB_FR be the fused result set of the feedback information, and FResult_Context[k] be a piece of semantic message of the fused result, then the fused result set TFB_FR can be expressed as:

$$TFB_FR = \{FResult_Context[k] | k=1,2,3,\ldots N\},$$

where, element FResult_Context[k] in the fused result set TFB_FR is the corresponding information of the element (InputSM[i]_SValue[j]) whose value is TRUE or FALSE of the status information set SM_SS in the information set SM_CC or SM_OC.

3.2 Integration Algorithm of the Feedback Information

In order to compare the redundant information fed back from multiple sensing channels and determine the final feedback information, the integration algorithm is designed as follows:

The reliability grade InputSM[i]_StableLevel of the channel whose value is TRUE or FALSE in state feedback information set SM_SS can be gotten respectively, and defining the summations for both situation as ST and SF respectively. If ST is bigger or equal to SF, this will indicate that the requested task has already been accomplished, and then the integrated result set TFB_FR can be expressed as:

$$TFB_FR = \{FResult_Context[k] | k=1,2,3,\ldots N\}$$
$$= \{SM_CC \cup SM_OC | InputSM[i]_SValue[j]$$
$$= TRUE,\ InputSM[i]_SValue[j] \in SM_SS\}$$

The element in the TFB_FR is the corresponding information InputSM[i]_CContext[j] in the information set SM_CC or InputSM[i]_OContext[j] in the information set SM_OC whose value is TRUE in status information set SM_SS .

If ST is smaller than SF, this will indicate that the robot has not finished the task, and then:

$$TFB_FR = \{FResult_Context[k] | k=1,2,3,\ldots N\}$$
$$= \{SM_CC \cup SM_OC | InputSM[i]_SValue[j]$$
$$= FALSE,\ InputSM[i]_SValue[j] \in SM_SS\}$$

The element in the TFB_FR is the corresponding information InputSM[i]_CContext[j] in the information set SM_CC or InputSM[i]_OContext[j] in the original information set SM_OC whose value is FALSE in status information set SM_SS .

Through this information processing method, the semantic integration of the feedback information is realized. The output intermediate level classify the information type of the information integration result TFB_FR, and according to the different processing efficiency to each type of information of the output channel, the pre-output information can be output to user by the channel with highest efficiency. Defining

OutputM[i] as some kind of information output channel, and OM as the set of the output channels, then:

$$OM=\{OutputM[i]|i=1,2,\ldots M\} = \{\text{video channel, speech channel, interface tips } \ldots\}$$

Defining OutputM[i]_InfoClass[j] as the information that channel OutputM[i] can output with the InfoClass[j] types information, and OM_C as the type set of the output information, then:

$$OM_C=\{OutputM[i]_InfoClass[j]|i=1,2,\ldots M, j=1,2,\ldots N\} = \{\text{image information, text information, digital signal, icon, speech } \ldots\}$$

Let OutputM[i]_InfoClass[j]_EfficiencyLevel be the efficiency level for channel OutputM[i] outputting InfoClass[j] type information, and OM_CE be the set of efficiency level for different channel outputting InfoClass[j] type information. By the way, efficiency level can be classified into 1, 2, 3, and so on, then:

$$OM_CE=\{OutputM[i]_InfoClass[j]\ _EfficiencyLevel|i=1,2,\ldots M, j=1,2,\ldots N\} = \{1, 2, 3,\ldots\},$$

where, the bigger of the level, the higher of the output efficiency for this kind of information.

Let OutputM[i]_OutputContext be the final context outputted by channel OutputM[i], then OutputM[i]_OutputContext is the information corresponding to the OutputM[i]_InfoClass[j]_EfficiencyLevel maximum value in the information integration result set(TFB_FR). If the TFB_FR is more than output channels, the messages are exported one after another according to the InputSM[i]_StableLevel of the sensing channel from large to small, and the information fed back from channel with a smaller InputSM[i]_StableLevel will be neglected.

4 Experiments

Experiments have been done on the "human–robot interaction platform" system shown in figure 2. The system consists of an interaction platform with vision, speech and button input channels and mobile robot named Hero. The result information of tasks executing of the robot is fed back to the user through the video, speech and graphics text tips of the information platform. The main interface of the interaction platform is shown in figure 3. After the user choose an interactive task, the dialog corresponding with the task begins to flash on the interface, and at the same time the interface communicate with the user through speech and information tip columns.

Interfaces(a) and (b)as shown in figure 3 show the process of choosing task of " lamp", the dialog(upper left corner) corresponding with task of lamp begins to flash on the interface, and background of the interface changes between figure (a) and figure (b) alternately. The message in the information column is changed to "You have chosen 'lamp' task, are you sure?", and meanwhile speech message also prompts current task to the user. After the user chooses "ok", message on the interaction interface is changed to the state as shown in figure 4.

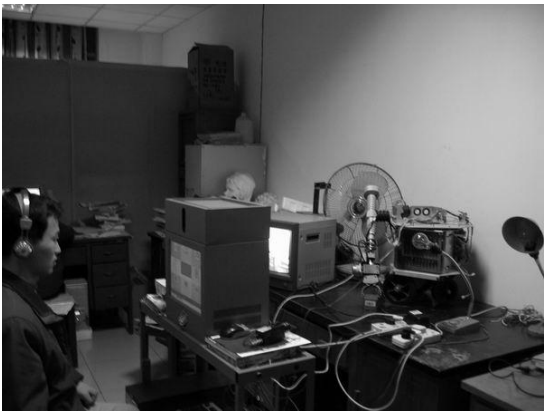
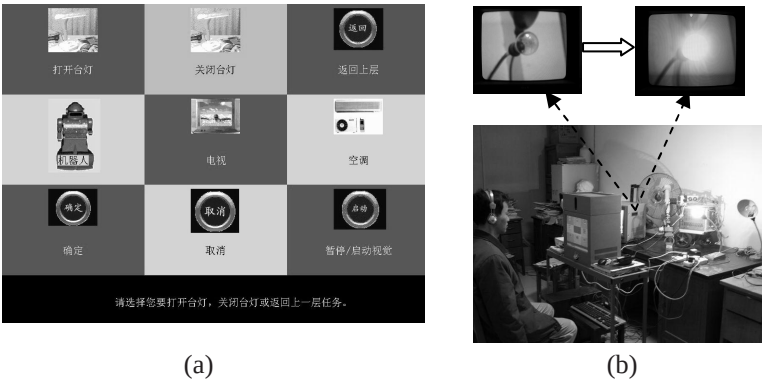


Fig. 2. Experiment environment for human-robot interaction

(a) (b)

Fig. 3. Information feedback and prompt in interaction interface



(a) (b)

Fig. 4. Information feedback results

When the state of human-robot interaction becomes to the next interface as shown in figure 4(a), the user choose the task “open lamp” , and confirm this selection, the robot executes the action of opening lamp, and the result is shown in figure 4(b). The user can obtain the result of executing tasks from video information, speech information and interface text tips feedback by CCD camera and monitor, meanwhile the user can feel the result of executing tasks from the change of environmental light intuitively. The reliability of visual channel in the feedback channels of this system is the highest and also the efficiency of feedback information of this channel is high, followed by the speech channel and feedback of the interface text tips.

5 Conclusions

It is very important that the results of tasks executed is feedback to the user in a reliable and intuitive way, especially the interaction process which is in remote or doesn’t facilitate to be observed directly, and the success or lost of tasks executed will be decided by feedback information of the tasks executing process and result in the system. And a kind of information feedback structure for human-robot interaction is presented, and expression and fusion method of feedback information is also proposed. According to the redundancy of multi-channel feedback information and complementarities among messages, abstract information fed back from multiple channels is integrated in semantic level. And according to characteristics of each feedback channel and the efficiency of each kind of channel feeding back different information, the integrated result information is classified and output finally. The ultimate output information is fed back to the user in an intuitive and visual channel. The cognitive strength of the user is reduced through this semantic information integration and feedback way, and mutual learning and adaptation in the process of human-robot interaction is enhanced. Meanwhile the information has the characteristic of redundancy and complementarity which is fed back from different angles using multiple channels, and a intuitive and accurate judgment about the state of tasks executing of the robot can be made, and the stability and efficiency of the human-robot process can also be improved.

Acknowledgements. This work was supported by Shanghai education committee for training excellent young teacher and the fund of Shanghai education committee (1001090607).

References

1. Hall, D.L., Llinas, J.: An Introduction to Multisensor Data Fusion. Proceedings of the IEEE 85(1), 6–23 (1997)
2. Rajeev, S.: Toward Multimodal Human–Computer Interface. Proceedings of the IEEE 86(5), 853–869 (1998)
3. Wu, L.Z., Sharon, L.O., Cohen, P.R.: Multimodal Integration-A Statistical View. IEEE Transactions on multimedia 1(4), 334–336 (1999)

4. Fritsch, J., Kleinehagenbrock, M., Lang, S., et al.: Multi-modal Anchoring for Human-Robot Interaction. Robotics and Autonomous Systems 43, 133–147 (2003)
5. Tao, P.J., Dong, S.H.: A Task-Oriented and Hierarchical Multimodal Integration Model and Its Corresponding Algorithm. Journal of Computer Research & Development 38(8), 966–971 (2001)

Multi-parameter Differential Pressure Flowmeter Nonlinear Calibration Based on SVM*

Jingqi Fu, Jing Li, and Ling Wang

Shanghai Key Laboratory of Power Station Automation Technology School of Mechatronics and Automation, Shanghai University Shanghai 200072, P.R. China

Abstract. In this paper, a novel modeling method of difference pressure mass flow measurement nonlinear correction based on Support Vector Machine is introduced. Support Vector Machine is a novel machine learning method, which identify the correction model definitely just according to the samples. Flow measurement nonlinear correction is a small sample problem. The result of the simulation for orifice flowmeter error correction based on SVM shows that this method can get a better effect.

Keywords: SVM, difference pressure flowmeter, data fusion.

1 Introduction

Mass flow is a very complex measurement. It is related to various parameters to a certain extent that hard to analysis in theory. Conventional methods can not establish a suitable model structure. Hardware compensation is difficult to achieve full compensation; the traditional modeling methods can not reach the requirements of high precision, such as linear least square, independent straight-line and so on; high order curve fitting is too complicated to realize.

With the development of computer technology, Data Fusion provides a new solution based on input-output data and expectation of output data, training the model can get the more accurate mass flow. At present, two common methods of data fusion are probability statistics and artificial neural network. Artificial neural network is widely used in the sensor calibration modeling because of good nonlinear approximation capability. However, several drawbacks, including local minimum, over learning, and the difficulties of choosing network structure, confine ANN's applications and development greatly. Support Vector Machine（SVM）is a new learning model. It is focused on a limited sample. Compared with traditional neural networks, SVM has many attractive characters; it has strictly theoretical foundation, good generalization capability, no local minima. Therefore, it provides a new modeling of multi-parameter differential pressure input-output nonlinear calibration by SVM[1]. The experimental results show SVM have advantages on the calibration accuracy and generalization ability compared with other correction models.

* This work was supported by National High Technology 863 Key Project(No. 2007AA04Z433563), Ministry of Science and Technology of China.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 896–903, 2008.
© Springer-Verlag Berlin Heidelberg 2008

2 Theory of Differential Pressure Mass Flow Measurement

Theory of differential pressure mass flow measurement is based on the fluid mechanical energy conversion principle. When fluid through the throttling element of the pipe, there will be static pressure difference between the upstream and downstream sides of throttling element. Differential pressure flowmeter works based on physical relationship between static pressure difference Δp and flow.

$$q_m = q_v \rho = \varepsilon \frac{\pi}{4} d^2 \frac{C}{\sqrt{1-\beta^4}} \sqrt{2\Delta p \rho} = K\sqrt{2\Delta p \rho} \tag{1}$$

where: q_v -volume flow, m^3/s; q_m -mass flow, kg/s; C-outflow coefficient (dimensionless); ε -expansion coefficient (dimensionless); β -diameter ratio, $\beta = d / D$; d-diameter of throttling element in the conditions of working, m; D-the diameter of upstream pipe in the condition of working, m; Δp -Differential pressure, Pa; ρ -density of upstream fluid, kg/m^3.

From formula (1) we know that flow measurement is related to various parameters to a certain extent. If ignoring the factor which affects the flow measurement slightly, differential pressure, pressure and temperature are still have very close connection to flow.

3 SVM Regression Algorithm

In SVM regression, the basic idea of SVM is firstly to map the input data into a high dimensional feature space via a nonlinear map. In the high dimensional feature space, linear decision function is constructed. Then SVM nonlinearly maps inner product of feature space to the original space via a kernel function.

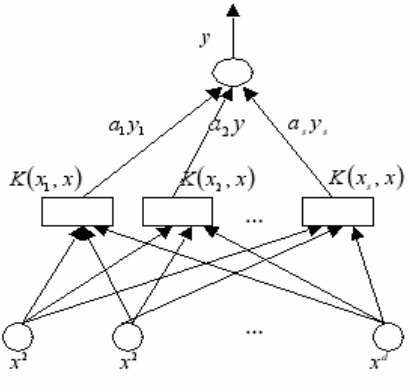


Fig. 1. SVM schematic diagram

It shows the nonlinear Support Vector Machine Vector diagram in Fig.1. $K(x,x_i)$ is the kernel function (kernel function is inner product function based on Mercer's condition). The decision function obtained by Support Vector Machine is similar to neural network formally, the output is linear combination of some intermediate nodes, and each node of intermediate layer corresponds to a support vector machine.

Different kernel functions form different SVM with different performances; kernel selection is the key problem when SVM is used to solve the regression problem. Here are several kernel functions used generally [2]:

Polynomial:

$$k(x,x^T)=(\langle x,x^T\rangle+1)^d \tag{2}$$

Gaussian Radial Basis Function (RBF):

$$k(x,x^T)=\exp(-\|x-x^T\|^2/(2\sigma^2)) \tag{3}$$

Multi-Layer Perception (Sigmoid Function):

$$k(x,x^T)=\tanh(\rho\langle x,x^T\rangle+b) \tag{4}$$

4 Modeling

The relationship between flow and pressure, temperature, differential pressure based on given data set $D=\{(u_1,q_1),(u_2,q_2),\cdots,(u_k,q_k)\}$ can be simply expressed as $q=f(u)$,Where $u_i \in R^3$ is 3-dimensional input vector, $q_i \in R$ (i=1,2,…k is the size of training data) is the output .The correction function is $z=f^{-1}(q)$, the training data set is: $D=\{(q_1,z_1),(q_2,z_2),\cdots,(q_k,z_k)\}$. Mapping the input space into a high dimensional feature space by nonlinear conversion $q \to \varphi(q)$ [2], that is

$$f^*(q)=\omega\varphi(q)+b \tag{5}$$

According to structure risk minimization principle, the optimization problems transform into

$$\min \frac{1}{2}\|\omega\|^2+c\sum_{i=1}^{k}(\xi_i+\xi_i^*) \tag{6}$$

formula (6) is called objective function. Where c is given constant, ξ_i^*, ξ_i are slack variables. By using ε-insensitive loss function $L(y,f(x,a))=L(|y-f(x,a)|_\varepsilon)$, so the constraint condition is $f^*(q_i)-z_i \le \xi_i^*+\varepsilon$, $z_i-f^*(q_i) \le \xi_i+\varepsilon$, $\xi_i^*,\xi_i \ge 0$. By introducing Lagrange function, the objective function transforms into

$$L(w,\xi_i,\xi_i^*)=\frac{1}{2}\|\omega\|^2+c\sum_{i=1}^{k}(\xi_i+\xi_i^*)-\sum_{i=1}^{k}a_i(\varepsilon+\xi_i+zi-f^*(q_i))-\sum_{i=1}^{k}a_i^*(\varepsilon+\xi_i^*+z_i-f^*(q_i))-\sum_{i=1}^{k}(r_i\xi_i+r_i^*\xi_i^*) \tag{7}$$

where $a_i,a_i^*,r_i,r_i^* \geq 0$. From the optimality conditions $\frac{\partial L}{\partial w}=0$, $\frac{\partial L}{\partial b}=0$, $\frac{\partial L}{\partial \xi_i}=0$, $\frac{\partial L}{\partial \xi_i^*}=0$, we obtain:

$$\omega=\sum_{i=1}^{k}(a_i-a_i^*)\varphi(q_i) \tag{8}$$

$a_i-a_i^*\neq 0$ is corresponding to support vector. We define kernel function $K(q,q_i)=\varphi(q_i)\cdot\varphi(q_j)$. By constraint condition and formula (8), optimization problem can be rewritten as

$$\max\omega(a_i,a_i^*)=\sum_{i=1}^{k}z_i(a_i-a_i^*)-\frac{1}{2}\sum_{i,j=1}^{k}(a_i-a_i^*)(a_j-a_j^*)K(q,q_i)-\sum_{i=1}^{k}(a_i+a_i^*)\varepsilon \tag{9}$$

subject to the constrains $\sum_{i=1}^{k}(a_i-a_i^*)=0, a_i,a_i^*\in[0,c],i=1,\cdots k$. Finally, the nonlinear regression function is obtained as

$$f^*(q)=\sum_{SVS}(a_i-a_i^*)k(q_i,q_i^*)+b \tag{10}$$

The key point of modeling is choosing error, kernel function and its parameters.

5 Simulation

We use the calibration data of GKF200-40 orifice flowmeter to verify the performance of nonlinear error correction based on SVM. The calibration data is shown in Table 1.

5.1 Normalization

Before training the model, data need to be normalized into the scale of (0, 1). Input and output normalization is based on formula (11) and formula (12):

$$\overline{x_i}=\frac{x_i-\min(x_i)}{\max(x_i)-\min(x_i)} \tag{11}$$

$$\overline{P_m}=\frac{0.9(P_m-\min(P_m))}{\max(P_m)-\min(P_m)}+0.05 \tag{12}$$

where $\overline{x_i}$ is the input value after being normalized, $\overline{x_i} \in (0,1)$; $\overline{p_m}$ is the output value after being normalized, $\overline{p_m} \in (0.05,0.95)$.

Table 1. Calibration Value

pressure	Difference pressure	temperature	Mass flow	pressure	Difference pressure	temperature	Mass flow
（MPa）	（KPa）	（℃）	（kg/h）	（MPa）	（KPa）	（℃）	（kg/h）
1.45	36.9	21.0	39573	1.64	31.5	18.0	28652
1.57	36.2	21.4	40124	1.62	32.7	17.9	31803
1.33	37.9	21.2	39693	1.25	39.0	17.6	40712
1.44	37.2	22.0	39753	1.42	25.9	13.9	33067
1.39	37.5	20.8	39646	1.26	37.3	17.4	39438
1.49	32.0	23.8	29886	1.38	38.4	16.5	42589
1.48	32.0	23.4	29400	1.55	26.0	15.1	29631
1.48	32.0	21.6	29651	1.32	37.3	15.2	28990
1.29	33.5	20.0	29261	1.47	36.4	15.7	39205
1.31	37.9	21.6	39245	1.49	37.1	15.5	40014
1.33	38.1	18.0	40905	1.46	37.1	15.7	40551
1.63	32.1	18.3	33499	1.58	36.7	15.3	40860

5.2 Results and Discussion

The curves of simulation based on SVM with different parameters of different kernel functions are shown below. The error curves of simulation based on Sigmoid Function are shown in Fig.2 with parameters $c=\infty$, $\rho=0.2$, $\varepsilon=0.00001$ and in Fig.3 with $c=2$, $\rho=2$, $\varepsilon=0.00001$.

The error curves of simulation based on SVM with Radial Basis Function is shown in Fig.4.

The simulated curve of model and real curve based on Radial Basis Function is shown in Fig.5, where point line is simulated curve and solid line is real curve.

The test results compared between Sigmoid Function and Radial Basis Function shows that Radial Basis Function has better performances when it is used to nonlinear correction. The simulation results based on Radial Basis Function coincide with the real values well, precision of estimated values is higher, and changing trend of system is tracked well. The results compared with other modeling methods BP algorithm are shown in Table 2. From Table 2 we can see that the model based on Radial Basis Function has good ability of regularization.

The simulation study shows that the methods of nonlinear error correction based on SVM have advantages as follows:

1. It is fit for nonlinear system identification because it processes nonlinear system only according to the input-output.
2. It only needs little known knowledge of nonlinear system.
3. The model based on SVM has good ability of generalization [1].

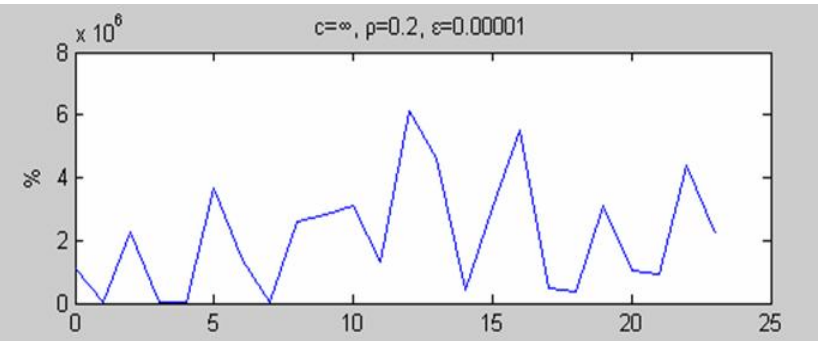


Fig. 2. The error curve with c=∞, ρ=0.2, ε=0.00001

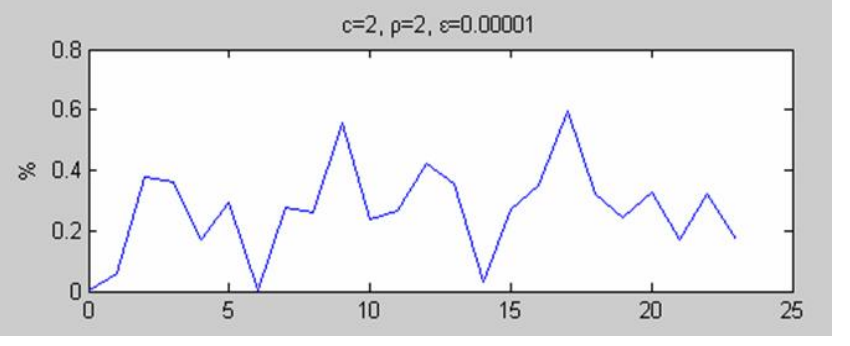


Fig. 3. The error curve with c=2, ρ=2, ε=0.00001

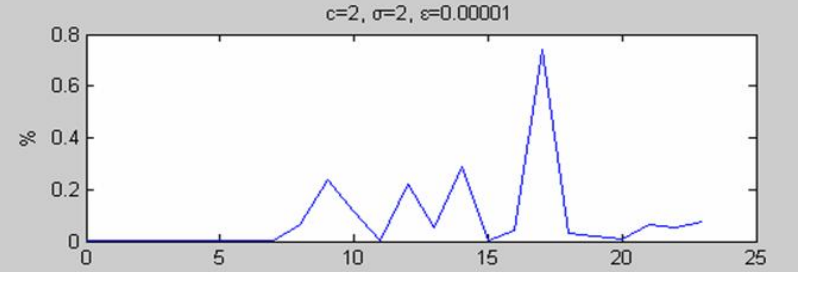


Fig. 4. The error curve with c=2, σ=2, ε=0.00001

At the same time, the simulation results based on Radial Basis Function shown in Fig.5 are not very well, the discrepancies between the real values and simulated values are not very ideal. It needs to do further study on choosing right parameters and kernel function.

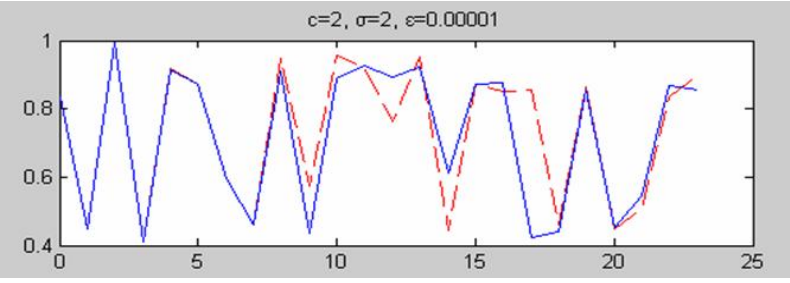


Fig. 5. The simulated curve and real curve based on Radial Basis Function

Table 2. Modeling Errors of Nonlinear Correction

Model	Function	parameter	MSE
BP	traingd	Epochs=5000,ir=0.01,goal=1e-4	0.020
SVM	Sigmoid Function	c=∞, ρ=0.2, ε=0.00001	2.6e12
		c=2, ρ=2, ε=0.00001	0.032
	RBF	c=∞, σ=0.2, ε=0.00001	0.013
		c=2, σ=2, ε=0.00001	0.011

6 Conclusions

Support vector machine (SVM) is a novel machine learning method. It is powerful for the nonlinear problem with small samples, high dimension, and local minima. In this paper, differential pressure nonlinear correction based on SVM was discussed and the corresponding simulation was implemented. The simulation results show that the method based on SVM is effective. However, to choose more suitable kernel function and its parameters of the algorithm, we need to do further study.

Acknowledgments. This research is financially supported by National Natural Science Foundation of China under grant 60774059, Key Project of Science & Technology Commission of Shanghai Municipality under Grant 061107031, Sunlight Plan Following Project of Shanghai Municipal Education Commission under grant 06GG10, and Shanghai Leading Academic Disciplines under grant T0103.

References

1. Nello, C., John, S.T.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Translated by li, G., Wang, M., Zeng, H. Tsinghua Publishing House of Electronics Industry, Beijing (2006)
2. Rong, H., Zhang, G.X., Zhang, C.F.: Application of Support Vector Machines to Nonlinear System Identification, April 4-8, pp. 501–507. IEEE press, Los Alamitos (2005)
3. Chen, W., Jie, Z., Chen, X., Chen, F., Cai, H.: Study on a SVM-based Data Fusion Method. Mechanical and electronic (8), 8–11 (2004)

4. Wei, Y.M.: Data Fusion Processing of Sensor Based on Neural Network. Journal of Data Acquisition & Processing 21 (Suppl.), 218–221 (2006)
5. Hsu, C.W., Chang, C.C., Lin, C.J.: A Practical Guide to Support Vector Classification (2008), http://www.csie.ntu.edu.tw/~cjlin
6. Document for MATLAB Interface of LIBSVM, http://www.csie.ntu.edu.tw/~cjlin/libsvm
7. Chang, C.C., Lin, C.J.: LIBSVM: a Library for Support Vector Machines (2007), http://www.csie.ntu.edu.tw/~cjlin/

Optimization Algorithm for Scalar Multiplication in the Elliptic Curve Cryptography over Prime Field

Yuanling Hao, Shiwei Ma, Guanghua Chen, Xiaoli Zhang, Hui Chen, and Weimin Zeng

School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, P.R. China
yuanling_hao@shu.edu.cn

Abstract. Elliptic curve cryptography (ECC) is a major public-key cryptosystem standard in data security systems due to its high strength of security. But its application was limited because of the constraint of its speed. Therefore, it is much necessary to develop efficient method for ECC implementation. In this paper, the efficient algorithm for scalar multiplication in ECC was discussed and the corresponding comparative results were given. An optimization technique based on sliding-window non-adjacent form (NAF-Window) algorithm for the high-speed operation of point doubling and point addition over prime fields in ECC system were proposed. The results of numerical tests and performance comparisons manifested that the proposed algorithm can remarkably improve the computational efficiency of scalar multiplication. Hence, it has practical significance for the implementation of ECC and is expected to be applied to data security.

Keywords: Elliptic curve cryptography, calar multiplication, Non-adjacent form, Sliding-window.

1 Introduction

In order to preserve and transmit information, the objectives of confidentiality, data integrity, data origin authentication, entity authentication and non-repudiation are expected in communication system. Therefore, cryptographic system is often employed in this field to ensure secure communications in the presence of malicious adversaries [1].

Elliptic curve cryptography (ECC) applies elliptic curve to cryptography, and has become an important branch of public-key cryptography system [2, 3]. Its security is based on the hardness of a different problem, namely the elliptic curve discrete logarithm problem (ECDLP) [4]. Because of short length of key and high strength of security, ECC has become the substitute for RSA and has been increasingly specified by accredited standards organizations (IEEEP1363, ANSI X9 and ISO / IEC, etc.) as a major public-key cryptosystem standards. Currently, a desired security level can be attained with significantly smaller key in elliptic curve systems than what is possible with their RSA counterparts. For example, it is generally accepted that a 160-bit elliptic curve key can provide the same level of security as a 1024-bit RSA key. Since short length of key makes ECC more suitable for resource-constrained environment. ECC

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 904–911, 2008.
© Springer-Verlag Berlin Heidelberg 2008

protocols in their security products, such as smart cards, PDA and high-bandwidth digital content protection (HDCP), etc.

Despite of the above-mentioned advantages of ECC, it still can not completely replace other public-key cryptosystems due to the constraint of its speed. Therefore, the research on the security of ECC and its efficient implementation method has become much necessary. In the implementation of ECC, scalar multiplication is not only the basic computing but also the most time-consuming operation [5]. So the operational efficiency of scalar multiplication directly determines the performance of ECC. In this work of study, the implementation of ECC over prime fields is discussed and ameliorated. An efficient algorithm of scalar multiplication for unfixed point is proposed and the corresponding numerical tests and performance comparative results are presented.

2 Elliptic Curve Arithmetic

The Weierstrass equation defined in the *G* domain is given by

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3 (a_i \in G) \tag{1}$$

It satisfies the condition of

$$F(X,Y,Z)(\frac{\partial F}{\partial X}, \frac{\partial F}{\partial Y}, \frac{\partial F}{\partial Z}) \neq (0,0,0) \tag{2}$$

Its solution contains an aggregate as

$$E\{(X,Y,Z) \in P^2(G) \big| F(X,Y,Z) = 0 \ \} \cup \ \{\infty\} \tag{3}$$

It is called elliptic curve *E* on *G* domain and can be given in affine coordinates by

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \ , a_i \in F \tag{4}$$

Let *p* be a prime number and Fp means the prime field. An elliptic curve *E* over Fp can be defined by

$$y^2 = x^3 + ax + b \tag{5}$$

where $a, b \in$ Fp , and $4a^3 + 27b^2 \neq 0$ in the Fp. Hence, the pair (x, y) , $x, y \in$ Fp , will be a point on the curve when (x, y) satisfies Eq.(5), and the point at infinity, denoted by ∞ , is said to be on the curve.

The rule of point addition and point doubling on elliptic curve *E* can be explained geometrically in Figure.2 [6]. Let $P(x_1, y_1)$ and $Q(x_2, y_2)$ be two distinct points on an elliptic curve *E*. The double *R*, of *P*, is defined by first drawing a tangent line to the elliptic curve at *P*, which intersects the elliptic curve at a second point, then *R* being the reflection of this point about the x-axis, depicted as Figure 2(b).

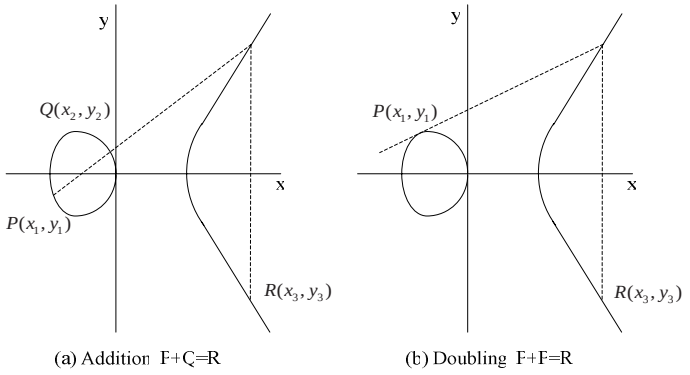


Fig. 1. Point addition and point doubling on elliptic curve

The algebraic formulas for these operational laws can be derived from the above geometric description, i.e

$$x_3 = \lambda^2 - x_1 - x_2 (\text{mod } p) \tag{6}$$

$$y_3 = \lambda(x_1 - x_3) - y_1 (\text{mod } p) \tag{7}$$

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1}, P \neq Q \\ \dfrac{3x_1^2 + a}{2y_1}, P = Q \end{cases} \tag{8}$$

where, the addition, subtraction, multiplication and inverse are arithmetic operations over Fp. Since the computing burden of addition and subtraction is negligible as compared with that of multiplication and inverse, only the operations of multiplication, square and the inverse will be considered in the following study.

The operation of $Q=kP$ is called point multiplication or scalar multiplication, where k is an integer and P is a point on an elliptic curve E defined over a field Fp , which dominates the execution time of elliptic curve cryptographic schemes. Since the implementations of the Elliptic Curve Digital Signature Algorithm (ECDSA) and the Elliptic Curve Diffie-Hellman (ECDH) in ECC mainly depends on the scalar multiplication kP, which consists of point addition and point doubling. The operational efficiency of kP is affected by the three main factors: the type of coordinate used, the algorithm of kP and the type of elliptic curve used.

3 Optimization of Point Addition and Point Doubling

In different coordinates, the forms of elliptic curve equation are also different and hence the efficiency of calculation will be different. Therefore, the selection of suitable

coordinates can avoid inverse operation which is the most time-consuming operation, and plays an important role in the improvement of efficiency.

Affine coordinate is one of the most commonly used coordinates in elliptic curve, in which the point addition and point doubling is given by Eq.(6)~(8). One can also use standard projective coordinate and Jacobian coordinate to avoid inverse operation. Table 1 gives the comparison of calculation by using these different coordinates, where *I*, *M* and *S* stand for inverse, multiplication and square respectively.

Table 1. Comparison of calculation in different coordinates

	point addition	point doubling
Affine coordinates	1+2M+S	1+2M+2S
Standard projective	12M+2S	7M+3S
Jacobian	12M+4S	4M+4S

From Table 1, one can see that the Jacobian coordinate is suitable to implement fast point doubling. The computational efficiency depends on ratio of inverse time to multiplication time, i.e. I/M . When $I > 12.4M$, this condition is easily established in most of the situations, the speed of point addition and point doubling is faster in Jacobian coordinate than that in affine coordinate. Researches also indicated that, when $I/M>10.6$, one should try to avoid using expressions which contain inverse operation [7]. This condition is also easily to be satisfied at present under the usual software and hardware environment. In addition, the number of calculations of point doubling is always much more than that of point addition (2~3 times generally) in the scalar multiplication operation. Therefore, the Jacobian coordinate supports high computational efficiency and is an optimization selection in ECC.

4 Optimization Algorithm of Scalar Multiplication

Generally, scalar multiplication *kP* in limited domain can use square-multiplication method which processes the scalar *k* to a binary encoding with point doubling of $\log_2(k)-1$ times. But the times of point addition are uncertainty for the variety of *k* [8]. In the circumstances of selecting *k* randomly, it can be assumed that the number of non-zero bits is only about half of the length of *k* in its binary expression, that is $\log_2(k)/2$. So the number of point addition is also $\log_2(k)/2$ and the lower bound of the times of point doubling will be $\log_2(k)-1$. In order to reduce the number of point addition, it is necessary to optimize the *kP* .

4.1 Non-adjacent Form

The computing efficiency of *kP* mainly depends on two factors: the length of $S(k)=(k_{l-1},...,k_0)$ and the number of nonzero elements of in $S(k)$. By using

non-adjacent form (NAF), the positive integer k may have the form of $k = \sum_{i=0}^{l-1} k_i 2^i$, where $k_i \in \{0, \pm 1\}$, $k_{l-1} \neq 0$, and no two consecutive digits k_i are non-zero, i.e. $k_i \times k_{i-1} = 0$. Hence, the number of non-zero elements in $S(k)$ can be reduced. By this way, the times of point addition can be reduced and thereby the speed of the computation of kP can be increased.

For each positive integer, the NAF is sole, such as

original: $1\,1\,1\,1\,0\,0\,1\,1\,0\,0\,0\,0\,0\,0\,1\,1\,1$

transform: $1\,0\,0\,0\,\bar{1}\,0\,1\,0\,\bar{1}\,0\,0\,0\,0\,0\,1\,0\,0\,\bar{1}$

It is based on the equation $2^{i+j-1} + 2^{i+j-2} + \ldots + 2^i = 2^{i+j} - 2^i$ to reduce the number of non-zero in the expansion of its scalar part. Because NAF of a positive integer is one bit longer at most compared with the length of a binary, and the average number of non-zero is only $\log_2(k)/3$ [9]. Therefore, NAF can enhance the speed about 11% more than that of the binary representation, in average.

4.2 Optimization Algorithm

Based on NAF method, we used sliding-window technology to further reduce the number of point addition. For elliptic curve and the point P on it, the proposed optimization algorithm is described as follows. Let the width of a sliding window be ω and all the possible values in width-ω window be m. One can compute the corresponding mP and preserve it in advance. When calculating kP, by searching the representation of k in term of NAF, i.e. $NAF(k)$ from left to right (from high bit to low bit), first lookup the window composed of continuous ω bits in which the first bit is non-zero, then find out the related mP value preserved and compute the corresponding point addition and point doubling. The detailed algorithms are given as follows.

Algorithm 1. Optimization algorithm of scalar multiplication for unfixed point

Input: $P(x,y) \in (Fp)$, positive integer k

Output: $Q = kP$

1. Computing NAF(k), $NAF(k)=(k_{n-1}, k_{n-2}, \ldots, k_1, k_0)_{NAF}$
2. Computing and preserving mP
3. $P_k \leftarrow P_{[k_{n-1}, k_{n-2}, \ldots, k_{n-\omega}]NAF}$, $i \leftarrow n\text{-}\omega$
4. While $i > \omega - 1$ do:
   - 4.1 searching the longest cluster of NAF, $k_{i-1}, \ldots, k_l$, satisfy $k_{i-1} = \ldots = k_l = 0$,
     $d \leftarrow i\text{-}l$, $i \leftarrow l$
   - 4.2 if $i > \omega - 1$ then:
     if $[k_{i-1} \ldots k_{i-\omega}]_{NAF} > 0$ then:
     $P_k \leftarrow 2^{d+\omega} \times P_k + P_{[k_{i-1}, \ldots, k_{i-\omega}]NAF}$

if $[k_{i-1}...k_{i-\omega}]_{NAF}<0$ then :

$P_k \leftarrow 2^{d+\omega} \times P_k - P_{[k_{i-1},...,k_{i-\omega}]NAF}$

$i \leftarrow i-\omega$

5. Computing : $P_k \leftarrow 2^i P_k + [k_{i-1}...k_0]_{NAF} P_k$

6. Return (P_k)

For the situation of fixed-point, *mP* can be computed off-line in advance (called pre-computation) and be used in the calculation of *kP*. The NAF(*k*) is divided into strings k_i and each one has the same length ω for $NAF(k) = k_{d-1} // ... // k_1 // k_0$. Because every k_i is non-adjacent, it can be denoted as an integer in the interval $[-1,1]$. The detailed algorithm is given as follows.

Algorithm 2. NAF-window algorithm for fixed point

Input : positive integer k , fixed point $P \in E(Fp)$

Output : kP

1. Pre-computing $P_i = 2^{wi} P$

2. Encoding $k = \sum_{i=0}^{[(l+1)/w]-1} k_i' 2^{wi}$, $k_i' \in [-2^{w-1}, 2^{w-1})$

3. Initializing $A \leftarrow \infty, B \leftarrow P$

4. Point multiplication : for $j=2^{w-1}$ to 1 , do

4.1 for i $\forall k_i' = j$, do : $B = B + P_i$

4.2 for i $\forall k_i' = -j$, do : $B = B - P_i$

4.3 $A \leftarrow A + B$

5. Return(A)

Obviously, the greater the ω, the more the points and storage space are required in Algorithm 4.2.2. In practice, the choice of ω should be based on the actual memory resources so as to ensure the universality of this algorithm.

4.3 Numerical Experiments and Results

In the proposed algorithm, the width of window affects its performance. If the width is too small, both the unnecessary judgments and the times of point addition will increase. But if it is too large, more computation in advance will be needed and more storage space will be required.

For unfixed point, the width of window depends on the value of *k*. After our calculation, when the bits of *k* is large (more than 100 bits), ω=4 is prefer, and when it is less than 100 bits, the election would be ω=3. For example, when ω=4, by adopting NAF, the potential window values are 1000, 1001, 1010, 100-1, 10-10, as well as the

corresponding negative values which can be easily calculated in elliptic curve. It is only expected to compute and storage 6*P*, 7*P*, 8*P*, 9*P* and 10*P* in the pre-computation. Other values can be obtained by their negation. Therefore, it only needs seven steps to complete the pre-computation.

In our experiment of numeric tests, the test platform is Intel P4 2.80GHz CPU, 512MB memory. The code of above algorithms was written in System C and using Visual C++ 6.0 as compiler. Elliptic curve domain parameters selected 192-bit over Fp. Table 2 gives the comparison in terms of speed and performance improvement of the binary representation, NAF and optimization algorithm. Table 3 gives the comparison of the computational complexity of point multiplication algorithms.

Table 2. Comparison of speed and performance

	Speed(times/sec)	Improvement (%)
Binary algorithm	1129.725	
NAF	1270.934	12.5
Optimized algorithm	1401.392	10.3

From these results, we can see that the times of point addition decrease by adopting the proposed optimization algorithm. Taken k=25722562047811804942 for an example, the times of point doubling and point addition in the optimization algorithm are 65 and 14 respectively with window width 4, while in binary algorithm they are 64 and 32 respectively. If we assume that point doubling and point addition spend roughly the same time, efficiency of the proposed optimized algorithm is higher than that of binary algorithm, and the improvement is $(64+32)-(65+14)/(64+32) \approx 0.1875$, that is, 18.75% increased.

Table 3. Comparison of the computational complexity

Algorithms	Complexity(average) $\log_2 k = m$		Storage requirements
	times of PD	times of PA	amount
Binary	m	m/2	2
W-rounding	$2^{w-1}+m$	$2^{w-1}+m/w$	2^{w}
NAF	m-1	(m-1)/3	2
Optimized algorithm..	m or m+1	$2^{w-2}+m/(w+1)$	2^{w-2}
Fixed point NAF-window	0	2 w-1+m/w	m/w

5 Conclusions

In the implementation of ECC, scalar multiplication is not only the basic computation but also the most time-consuming operation. Its operational efficiency directly determines the performance of ECC. The proposed optimization algorithm employed

sliding-window based NAF technique to reduce the times of computation of scalar multiplication over prime fields for unfixed point in elliptic curve. Theoretic analyses and numerical tests manifested that this algorithm can remarkably enhance the computing efficiency of scalar multiplication compared with other traditional algorithms and therefore have practical significance for the implementation of ECC.

Acknowledgements. The research is financially supported by Shanghai Education Commission Project (07ZZ03).

References

1. Lauter, K.: The Advantages of Elliptic Curve Cryptography for Wireless Security. J. IEEE Wireless Communications 11(1), 62–67 (2004)
2. Koblitz, N.: Elliptic Curve Cryptosystems. J. Math. Compute. 48, 203–209 (1987)
3. Menezes, A.: Elliptic Curve Public Key Cryptosystems. Kluwer Academic Publishers, Dordrecht (1993)
4. Menezes, A., Koblitz, N., Vanstone, S.: The State of Elliptic Curve Cryptography. Des. Cryp. 19, 173–193 (2000)
5. Certicom Corp. for ECC, http://www.certicom.com/
6. Satoh, T.: The Canonical Lift of an Ordinary Elliptic Curve Over a Finite Field and its Point Counting. J. Math. Soc. 15, 247–270 (2000)
7. Cohen, H., Miyaji, A., Ono, T.: Efficient Elliptic Curve Exponentiation Using Mixed Coordinates. In: Ohta, K., Pei, D. (eds.) ASIACRYPT 1998. LNCS, vol. 1514, pp. 51–65. Springer, Heidelberg (1998)
8. Tenca, A.F., Savas, E., Koc, C.K.: A Design Framework for Scalable and Unified Multipliers in GF(p). J. Com. Rese. 13, 68–83 (2004)
9. Egecioglu, O., Koc, C.K.: Fast Modular Exponentiation. In: Proceeding of 1990 Bilkent International Conference on New Trends in Communication, Control, and Signal Processing, pp. 188–198. Elsevier, Amsterdam (1990)

A Vague Sets Based Hierarchical Synthetic Evaluation Algorithm for Health Condition Assessment*

Hua Bai[1] and Lixin Gao[2]

[1] School of Mechatronics Engineering, Harbin Institute of Technology, Harbin, China
baihua@hit.edu.cn
[2] School of Municipal and Environmental Engineering, Harbin Institute of Technology, Harbin, China
gaolixin@hit.edu.cn

Abstract. The hierarchical synthetic evaluation algorithm for health condition assessment of buried water pipeline is studied. Usually, the deterioration of buried pipeline is influenced by many uncertainty factors, such as material, surrounding environment, etc. It is essential to properly represent and use multi-source information to conduct rational health condition evaluation analysis. By the use of a hierarchical framework, a vague sets based synthetic evaluation algorithm has been developed for supporting such evaluation analysis, the kernel of which organizes data, measures similarity and evaluates the result on the basis of the vague sets. Compared with the existing fuzzy comprehensive evaluation method, this method is more powerful and efficient to deal with conflicting information as well as partial ignorance.

Keywords: vague sets, hierarchical synthetic evaluation algorithm, health condition, multi-source information fusion.

1 Introduction

With the increasingly large number of leaks and failure of mains occurred in portable water distribution networks, more and more concern focus on the study of how to translate pipe inspection results into corresponding health condition rating.

The deterioration of pipe is a physical manifestation of the ageing process, which influenced by many factors, such as the pipe materials, stability and composition of surrounding soils, car load-carrying capacity of road surface, internal flow rate and aggressiveness of transported water, etc. The pipe deterioration manifests the changes in internal and external surfaces, misalignment and displacement of joints, formation of corrosion pits, cracks and spalling of mortar, etc. Usually, one distress indicator could describe one kind of change caused by deterioration. Each of these distress indicators contributes differently to the condition rating of a given segment of a pipe. The physico-chemical and microbiological processes that govern deterioration are often not understood well enough to merit an adequate model based on mechanics or electro-chemistry

* This work is supported by the National Science Foundation of China (NSFC) under grant No. 60772077.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 912–920, 2008.
© Springer-Verlag Berlin Heidelberg 2008

or microbiology. Consequently there is a need to rely on the cumulative experience of experts in encoding of distress indicators into condition rating. This process, however, makes the assessment of pipe condition rating a challenging task and requires a large body of data as well as experts' judgment.

To identify different kinds of distress indicators, some nondestructive technologies (based on ultrasonic, eddy current, etc.) have recently been developed[1],[2] to inspect and monitor water pipelines. Typically, a given sensor or technology provides data about a specific aspect of a distress indicator with a degree of accuracy. Condition rating estimated from any isolated indicators is generally imprecise due to the limitation of the knowledge and understanding about the process, and sometimes even leads to a conflicting or contradicting assessment depending on the outcome of the measurements or assessment methods used. Therefore, a comprehensive and credible condition rating evaluation requires fusion of data obtained from several sources (sensors or distress indicators). Thus, we can obtain more accurate and credible fusion result than the result obtained by using the sensors solely.

In this paper, a data fusion approach which has the advantage of processing uncertain, inaccurate and conflicting information is proposed. Distress indicators are described as vague sets, arranged in a hierarchical tree-like structure, to evaluate health condition of water pipelines. The concept of vague sets is introduced in section 2. In section 3, necessary basics are briefly described about the vague sets based hierarchical synthetic evaluation algorithm, and the framework of the proposed evaluation analysis model. An example of pipe condition assessment is provided in section 4 to demonstrate the proposed approach. Finally, discussion and conclusions are provided in section 5.

2 Vague Sets

A vague sets is a set of objects, each of which has a grade-of membership which is characterized by a truth-membership function and a false-membership function. To some extent the vague sets is a special fuzzy sets which has a grade-of membership whose value is a continuous subinterval of [0, 1]. But compared to the exist fuzzy sets, it is more powerful in the describing and processing of uncertain and inaccurate information, in particular of conflicting information.[5]

A fuzzy sets F, as introduced by Zadeh[3], is a class of objects X along with a grade of membership function. This membership function $\mu_F(x), x \in X$, assigns to each object a grade of membership ranging between zero and one. Zadeh assigns each object a single value. This single value combines the evidence for $x \in X$ and the evidence against $x \in X$, without indicating how much there is of each. The single number tells us nothing about its accuracy. So, this technique do not rationally account for confirmatory and/or contradictory information.

Let X be a space of objects, with a generic element of X denoted by x. A Vague Sets V in X is characterized by a truth-membership function t_v and a false-membership function f_v. $t_v(x)$is a lower bound on the grade of membership of x derived from the evidence for x, and $f_v(x)$is a lower bound on the negation of x derived from the evi-

dence against x. $t_v(x)$and $f_v(x)$both associate a real number in the interval [0,1] with each object in X, where $t_v(x)+ f_v(x)\leq 1$. That is $t_v : X \to [0,1]$, $f_v : X \to [0,1]$.

This approach bounds the grade of membership of x to a subinterval $[t_v(x), 1- f_v(x)]$ of [0,1]. In other words, the exact grade of membership $\mu_v(x)$ of x may be unknown, but is bounded by $t_v(x) \leq \mu_v(x) \leq 1- f_v(x)$, where $t_v(x)+ f_v(x)\leq 1$.

For example, the health condition of a pipe can be divide into three states, i.e. H = {*good*, *fair*, *bad*}, denoted as $H = \{H_1, H_2, H_3\}$. The pipe condition belongs to each state can be described as a vague sets. Assume that the information obtained is written as $[t_{H_1}, 1-f_{H_1}]=[0.5, 0.8]$, it describes $t_{H_1}=0.5$, and $1- f_{H_1}=0.8$. That is to say, based on the obtained information, the grade of membership for H_1 is 0.5, against H_1 is 0.2, and 0.3 is unassigned because of the uncertain and inaccurate information.

3 Basics about the Hierarchical Synthetic Evaluation Model

3.1 Framework of HSE Model

The hierarchical synthetic evaluation (HSE) model is a generic framework to aggregate and to handle various bodies of evidence in a recursive manner. In a HSE framework, the attribute at a higher-level is evaluated based on the assessment of its associated lower-level factors as illustrated in Figure1.

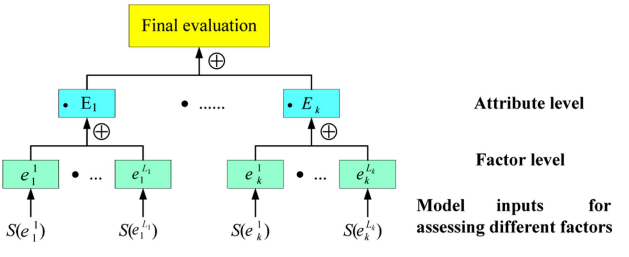


Fig. 1. Framework of the proposed HSE model

The condition rating of the k^{th} attribute, E_k, is evaluated based on a number of factors, which can be directly observed or estimated. The evaluation of an attribute E_k with contributory factors e_k^i (i = 1, 2, …, L_k) is given by,

$$E_k = e_k^1 \oplus e_k^2 \oplus \ldots \oplus e_k^{L_k} \tag{1}$$

where L_k denotes the number of factors that contribute to the k^{th} attribute, and $\oplus$ is a combination operator. Assume the condition states of a pipe is described as,

$$H = \{H_1, H_2, \ldots, H_j, \ldots, H_n\}; \quad j = 1, 2, \ldots, n \tag{2}$$

where n is the number of possible condition states, H_j represents the j^{th} condition state, and H_1 and H_n are the best and the worst possible condition states, respectively.

Through the inspection result, an expert may not always be 100% sure that the condition state is exactly confined to only one condition states. In most instances, condition rating will be confined to two or no more than three contiguous condition states with a total degree of confidence equal or smaller than 100%.

The condition assessment(S) for a given factor could be obtain from the inspection directly and written as,

$$S(e_k^i) = \{(H_j / (t_{H_j,i}, 1 - f_{H_j,i})), j = 1, \ldots, n\}; i = 1, 2, \ldots, L_k \quad (3)$$

where $t_{H_j,i} + f_{H_j,i} \le 1$. Thus, the factor, e_k^i, can be assessed for grade H_j with a degree of $t_{H_j,i}$, and against grade H_j with a degree of $f_{H_j,i}$. The confidence of an assessment is complete if $t_{H_j,i} + f_{H_j,i} = 1$, and is incomplete if $t_{H_j,i} + f_{H_j,i} < 1$.

For example, as shown in table1, two contributory factors contribute to evaluate the 4th attribute of a pipe, joint condition(E_4). Each of the factors consists of three condition states, namely, *good*, *fair* and *bad*, which can be written as,

$$H = \{good\ (H_1), fair\ (H_2), bad\ (H_3)\} \quad (4)$$

The condition assessment of factor e_4^1 is 40%for *good condition* with a confidence of 80%, that is to say, it is 40% for *good condition* and 40% against *good condition;* and it is 60% for *fair condition* with a confidence of 90%; and it is not in *bad condition* with a confidence of 100%. It is the same for the factor (e_4^2).

In practice, not all the factors have the same importance towards the assessment of an attribute. In addition, the data collected from different sensors may be erroneous (less reliable) or expert judgment may have different levels of credibility. All these influencing factors are lumped together into a parameter, λ_k^i, which is a normalized relative to the weight of evidence for a factor, e_k^i, towards the evaluation of an attribute, E_k. Therefore, the weight matrix for an attribute, E_k, can be written as,

$$\lambda_k = \{\lambda_k^1, \ldots, \lambda_k^i, \ldots, \lambda_k^{L_k}\} \quad \text{where} \quad 0 \le \lambda_k^i \le 1 \quad (5)$$

3.2 Summary of the HSE Algorithm

Combining the i factors that contribute to the k^{th} attribute with its weight, we could obtain the condition rating for attribute E_k as follows,

$$S(E_k) = \lambda_k \circ \begin{bmatrix} S(e_k^1) \\ \vdots \\ S(e_k^i) \\ \vdots \\ S(e_K^{L_k}) \end{bmatrix} = [\lambda_k^1 \cdots \lambda_k^i \cdots \lambda_k^{L_k}] \circ \begin{bmatrix} (t_{H_1,1}, 1 - f_{H_1,1}) & (t_{H_2,1}, 1 - f_{H_2,1}) & \cdots & (t_{H_n,1}, 1 - f_{H_n,1}) \\ \vdots & \vdots & \vdots & \vdots \\ (t_{H_1,i}, 1 - f_{H_1,i}) & (t_{H_2,i}, 1 - f_{H_2,i}) & \cdots & (t_{H_n,i}, 1 - f_{H_n,i}) \\ \vdots & \vdots & \vdots & \vdots \\ (t_{H_1,L_k}, 1 - f_{H_1,L_k}) & (t_{H_2,L_k}, 1 - f_{H_2,L_k}) & \cdots & (t_{H_n,L_k}, 1 - f_{H_n,L_k}) \end{bmatrix} \quad (6)$$

Two vague sets operators $\cdot$ and $\oplus$ which defined in [5] were used in this combination process.

$$k\cdot(t_v,1-f_v)=[k\cdot t_v,k(1-f_v)]\quad k\in[0,1]$$
$$(t_{v1},1-f_{v1})\oplus(t_{v2},1-f_{v2})=[t_{v1}\oplus t_{v2},(1-f_{v1})\oplus(1-f_{v2})]=[\min\{1,t_{v1}+t_{v2}\},\min\{1,(1-f_{v1})+(1-f_{v2})\}] \tag{7}$$

Thus, the condition rating for attribute E_k could be written as:

$$\begin{aligned} S(E_k) &= [\min\{1,\sum_{i=1}^{L_k}\lambda_k^i\cdot t_{H_j,i}\},\min\{1,\sum_{i=1}^{L_k}\lambda_k^i(1-f_{H_j,i})\}] \\ &= \{H_j/(t_{E_{k,j}},1-f_{E_{k,j}})\} \qquad Where \quad j=1,2,\cdots,n \end{aligned} \tag{8}$$

For each condition state H_j, the ideal condition is that it is 100% for condition H_j with a confidence of 100%, which written in a vague sets as (1,1). Thus, for the k^{th} attribute, its ideal condition could be written as:

$$S^*(E_k)=\{H_j/(1,1)\},\qquad j=1,2,\ldots,n \tag{9}$$

Some method has been proposed to measure the degree of similarity between vague sets. The distance between vague sets is an effective way to show the degree of their similarity. The less the distance is, the more the degree of similarity will be. Thus, based on the algorithm proposed in [5], we could calculate the distance between the combined $S(E_k)$ and its ideal $S^*(E_k)$which is written as T(S, S^*), as shown in(10), to determine the corresponding evaluation result.

$$T_j(S(E_k),S^*(E_k))=\left\{H_j/\tfrac{1}{2\times L_k}\left(\left|1-t_{E_{k,j}}\right|+\left|1-f_{E_{k,j}}\right|\right)\right\},\quad where\quad j=1,2,\cdots,n \tag{10}$$

4 Application for Pipeline Health Condition Evaluation

The following section illustrate the application of fusing multi-source data using the HSE model to evaluate a lined cast iron pipe condition from its four major attributes, which could be further evaluated with contributory factors shown in table1.

First, select the attribute, *internal surface condition* of a lined cast iron pipe, to serve as example illustrating the combination formulae discussed in section 3. The *internal surface condition* (k=1) depends on three contributory factors, namely, *spalling of the cement lining*, *internal corrosion pit depth*, and *tuberculation*. The aggregation of factors for the attribute, *internal surface condition* (E_1) is given by:

$$Internal\ surface\ condition\ (E_1)=Spalling(e_1^1)\oplus Pit\ depth(e_1^2)\oplus Tuberculation(e_1^3) \tag{11}$$

The condition rating $S(e_k^i)$ for each factor is provided in vague sets as indicated in Table1. For the first factor in the first attribute, i.e., k = 1 and i = 1, e_1^1 corresponds to the assessment that "cement lining spalling condition is 50% for *bad* (*B*) with a degree of confidence of 70% and 50%for *fair(F)* with a degree of confidence of 90%", which implies that there is ignorance or epistemic uncertainty information on assessing factor e_1^1 . Thus, the assessment $S(e_1^1)$ is written as {*G*/(0,1), *F*/(0.5,0.6),

B/(0.5,0.8)}, Similarly, $S(e_1^2)$ is {*G*/(0.5,0.7), *F*/(0.5,0.6), *B*/(0,1)} and $S(e_1^3)$ is {*G*/(0.6,0.8), *F*/(0.4,0.5), *B*/(0,0.1)}.

Table 1. Summary of attributes and the contributory factors that define the overall condition rating of a lined cast iron pipe

Attributes, *k*	**Contributory factors**, *i*	**Weight**		**Condition ratings of distress indicators**, $S(e_k^i)$	
		γ_k	λ_k^i	Vague sets	Fuzzy sets
Internal surface (E_1)	Cement lining spalling (e_1^1)	0.3	0.4	*G*/(0,1),F (0.5,0.6,),*B*/(0.5,0.8)	G/0, F/0.5 ,B/0.5
	Internal corrosion pit depth (e_1^2)		0.4	*G*/(0.5,0.7),F /(0.5,0.6),B/(0,1)	G/0.5, F/0.5 ,B/0
	Tuberculation (e_1^3)		0.2	*G*/(0.6,0.8), F/(0.4,0.5),*B*/(0,1)	G/0.6, F/0.4 ,B/0
External Pipe barrel (E_2)	External pit depth (e_2^1)	0.3	0.28	*G*/(0,1), F/(0.6,0.8),*B*/(0.4,0.7)	G/0, F/0.6 ,B/0.4
	Graphitization aerial extent (e_2^2)		0.28	*G/(0.4,0.8)*,*F*/(0.5,0.7), B/(0,1)	G/0.4, F/0.5 ,B/0
	Crack type (e_2^3)		0.16	*G*/(0.8,0.9),*F*/(0.2,0.4),*B*/(0,1)	G/0.8, F/0.2 ,B/0
	Crack width (e_2^4)		0.28	*G*/(0,1),*F*/(0.6,0.8),*B*/(0.4,0.5)	G/0, F/0.6 ,B/0.4
External coating(E_3)	Crack/tear (e_3^1)	0.3		*G*/(0.4,0.6),*F*/(0.6,0.7),B/(0,1)	G/0.4, F/0.6 ,B/0
Joint (E_4)	Change in alignment (e_4^1)	0.1	0.65	*G*/(0.4,0.6),*F*/(0.6,0.7),*B*/(0,1)	G/0.4, F/0.6 ,B/0
	Joint displacemen (e_4^2)		0.35	G /(0,1),*F*/(0.7,0.9),*B*/(0.3,0.4)	G/0, F/0.7 ,B/0.3

The assessment for the related factors are defined as {*good*(*G*)/β, *fair*(*F*)/β, *bad*(*B*)/β}, where β represents the grade of true/false membership for a vague sets or the grade of membership for a fuzzy sets.

The weights (importance and credibility of evidence) for each contributory factor to the attribute E_1 are assigned using experts' judgment. The weight matrix can be written as,

$$\lambda_1 = \{\lambda_1^1, \lambda_1^2, \lambda_1^3\}^T = \{0.4, 0.4, 0.2\}^T \tag{12}$$

Using the recursive combination formulae(8), the combined condition rating for attribute E_1 can be determined as follows.

$$S(E_1) = [\min\{1, \sum_{i=1}^{3} \lambda_1^i \cdot t_{H_j,i}\}, \min\{1, \sum_{i=1}^{3} \lambda_1^i (1 - f_{H_j,i})\}]$$
$$= \{H_1/(t_{E_{1,1}}, 1 - f_{E_{1,1}}), H_2/(t_{E_{1,2}}, 1 - f_{E_{1,2}}), H_3/(t_{E_{1,3}}, 1 - f_{E_{1,3}})\} \quad (13)$$

Here, $t_{E_{1,1}}$ =min{1,(0.4×0+0.4×0.5+0.2×0.6)}=0.32

$1 - f_{E_{1,1}}$ =min{1,(0.4×1+0.4×0.7+0.2×0.8)}=0.84

$t_{E_{1,2}}$ =min{1,(0.4×0.5+0.4×0.5+0.2×0.4)}=0.48

$1 - f_{E_{1,2}}$ =min{1,(0.4×0.6+0.4×0.6+0.2×0.5)}=0.58

$t_{E_{1,3}}$ =min{1,(0.4×0.5+0.4×0+0.2×0)}=0.2

$1 - f_{E_{1,3}}$ =min{1,(0.4×0.8+0.4×1+0.2×1)}=0.92

So the combined condition rating for attribute E_1 is {G/(0.32,0.84), F/(0.48,0.58), B/(0.2,0.92)}. By using(10), we could get the distance between the $S(E_1)$ and $S^*(E_1)$ is

$$T_1[S(E_1), S^*(E_1)] = \frac{1}{2\times 3}(1 - 0.32 + 1 - 0.84) = 0.14, \; T_2[S(E_1), S^*(E_1)] = \frac{1}{2\times 3}(1 - 0.48 + 1 - 0.58) = 0.157$$

$$T_3[S(E_1), S^*(E_1)] = \frac{1}{2\times 3}(1 - 0.2 + 1 - 0.92) = 0.15$$

Therefore, based on the least distance, the final condition rating for the attribute *internal surface condition* (E_1) is *good*. However, it is important to note that the degrees of confidence for the conditions *fair* and *bad* are not too different from the good. Therefore, these results are not conclusive and there is the need to involve other attributes for a more reliable decision.

Similarly, contributions of the remaining three attributes such as *external pipe barrel*, *external coating* and *joint conditions* are determined. All four attributes are subsequently combined to obtain the overall condition rating.

As depicted in Figure 1, the four attributes can be written in terms of the factors, as $E_1 = \{e_1^1 \quad e_1^2 \quad e_1^3\}$; $E_2 = \{e_2^1 \quad e_2^2 \quad e_2^3 \quad e_2^4\}$; $E_3 = \{e_3^1\}$; and $E_4 = \{e_4^1 \quad e_4^2\}$. Similarly, the weight vectors for each factor are defined as $\lambda_1 = \{\lambda_1^1, \lambda_1^2, \lambda_1^3\}$; $\lambda_2 = \{\lambda_2^1, \lambda_2^2, \lambda_2^3, \lambda_2^4\}$; $\lambda_3 = \{\lambda_3^1\}$; and $\lambda_4 = \{\lambda_4^1, \lambda_4^2\}$. Table1 summarizes the attributes and the contributory factors that define the overall pipe condition rating of a lined cast iron pipe. Each attribute is evaluated using the same aggregation procedure described in above. The overall condition assessment written in matrix form is

$$\begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{bmatrix} = \begin{bmatrix} (0.32,0.84) & (0.48,0.58) & (0.2,0.92) \\ (0.24,0.92) & (0.51,0.71) & (0.22,0.78) \\ (0.4,0.6) & (0.6,0.7) & (0,1) \\ (0.26,0.61) & (0.64,0.77) & (0.11,0.79) \end{bmatrix} \quad (14)$$

The matrix on the right-hand side of the above equation represents the combined assessment for each attribute (E_k). The columns of the matrix represent the degree correspond to condition states *good*, *fair* and *bad*. The contribution of each of these attributes to the overall condition rating can be described by the use of importance weights, which are assigned based on experts' judgment. We select the importance weight matrix as γ_k = {0.3, 0.3, 0.3, 0.1} for the lined cast iron pipe. Using recursive

combination formulae in Eqn.8, we could combine the four attributes to obtain the overall condition rating (*E*) of a lined cast iron pipe as shown below

$$S(E)=[\gamma_1 \quad \gamma_2 \quad \gamma_3 \quad \gamma_4]\circ \quad [S(E_1)S(E_2)S(E_3)S(E_4)]^T=\{G/(0.31,0.77) \quad F/(0.54,0.67) \quad B/(0.14,0.88)\} \tag{15}$$

And the distance between the *S(E)* and the $S^*(E)$ is $T_1[S(E),S^*(E)]=0.114$, $T_2[S(E),S^*(E)]=0.09$, $T_3[S(E),S^*(E)]=0.12$. Therefore, the most likely condition rating of this lined cast iron pipe is *fair*. It is also noted that the combined overall condition rating is very distinct from the condition rating based on one individual attribute or category.

5 Discuss and Conclusion

To light the advantage of the vague sets based HSE model, we define the same γ_k, λ_k^i, *H*, and the same value on the grade of membership for each factor, then combine the total factors and attributes by a fuzzy sets based HSE model. Select one of the evaluation operation models, as shown below, then the comprehensive evaluation result can be calculated by $S(E_k)=\lambda\circ S(e_k^i)$ *and* $S(E)=\gamma\circ S(E_k)$. Finally, use maximum membership principle to determine the result. Where the operation models are:

(1): $M(\wedge,\vee)$: $S_j=\bigvee_{i=1}^{m}(\lambda_i\wedge s_{i,j})$ (2): $M(\bullet,\vee):S_j=\bigvee_{i=1}^{m}\left(\lambda_i\bullet s_{i,j}\right)$ (3): $M(\wedge,\oplus):S_j=\bigoplus_{i=1}^{m}(\lambda_i\wedge s_{i,j})=\sum_{i=1}^{m}(\lambda_i\wedge s_{i,j})$

The comprehensive evaluation results can be calculated as follows respectively

(1): S=(G/0.3, F/0.3, B/0.3) , (2): S=(G/0.24, F/0.33, B/0.16),(3): S=(G/1, F/1, B/0.7)

To the data in this example, model (1) and (3) are poor to determine the right result by using the maximum membership principle. It is the weakness of a fuzzy sets that can't effectively fuse the incomplete or conflicting data from different sources.

AS presented in this paper, condition assessment of water pipeline reflects an aggregate state of its health, which requires a rational, repeatable, and transparent approach. Compared to the existing fuzzy approach, a vague sets based hierarchical synthetic evaluation (HSE) model is more effective to combine subjective, imprecise and incomplete information, and even conflicting data.

The above numerical analysis deals with a lined cast iron pipe, which demonstrated the implementation process of the new vague sets based HSE approach on a step-by-step basis. Due to its open architecture and a firm mathematic foundation, it can be modified at any level of hierarchical structure without changing the recursive combination algorithm. Although the example is taken for a buried water pipeline health condition evaluation, it is clear that the vague sets based HSE algorithm can be applied to general health condition evaluation problems with or without uncertainty.

References

1. Buehrer, D.J.: Vague logic: A First-Order Extension of Interval Probability Theory. In: Proceedings of IEEE Workshop Imprecise and Approximate Computation, pp. 83–87 (1992)
2. Duran, O., Althoefer, K., Seneviratne, L.D.: State of the art in Sensor Technologies for Sewer Inspection. IEEE Sensors Journal 2(2), 73–81 (2002)
3. Zadeh, L.A.: Fuzzy sets. Information control 8, 338–353 (1965)

4. Gau, W.L., Buehrer, D.J.: Vague sets. IEEE Transactions on Systems, Man and Cybernetics 23(2), 610–614 (1993)
5. Li, F., Xu, Z.Y.: Measures of Similarity Between Vague sets. Journal of software 12(6), 922–927 (2001)
6. Xu, Z.S.: Uncertain Multiple Attribute Decision Making: Method and Applications. Tsinghua University Press, Beijing (2004)
7. He, Y.: Multi-sensor Information Fusion with Applications. Electronic Industry Press, Beijing (2000)

Application of Fuzzy Classification in Bankruptcy Prediction

Zijiang Yang[1] and Guojun Gan[2]

[1] York University
zyang@mathstat.yorku.ca
[2] York University
gjgan@mathstat.yorku.ca

Abstract. Classification refers to a set of methods that predict the class of an object from attributes or features describing the object. In this paper we present a fuzzy classification algorithm to predict bankruptcy. Our classification algorithm is modified from a subspace clustering algorithm called fuzzy subspace clustering (FSC). As our algorithm associates each feature of a class with a fuzzy membership, feature selection is not necessary. Our experiments show that the classification results produced by our algorithm can translate into large financial and other benefits to organizations through such activities as credit approval, and loan portfolio and security management.

Keywords: Bankruptcy prediction, Fuzzy subspace clustering, Data mining, Classification.

1 Introduction

Bankruptcy prediction has always attracted significant global attention. Financial institutions, in particular, are interested in an effort to reduce the level of risk in their investments. Given the growing importance of the bankruptcy prediction, there have been many attempts to model the prediction of business failure.

The first publication appeared in 1966 and was authored by W. Beaver [2]who created a univariate discriminant model using financial ratios selected by a dichotomous classification test. Since then bankruptcy prediction models have evolved to use both statistical analysis and data mining techniques to refine the decision support tools and improve decision making. In addition to discriminant analysis, traditional statistical methods include regression, logistic models, factor analysis, etc. More recent data mining techniques include decision trees, neural networks (NNs), fuzzy logic, genetic algorithms (GA) and support vector machines (SVM) among others. The statistical applications, although enhanced over time, were restricted by the rigorous assumptions of traditional statistics such as: linearity, normality, independence among predictor variables and pre-existing functional form relating the criterion variable and the predictor variable.

Balcean and Ooghe [1] presented an overview of classic statistical methods for predicting business failure developed thus far and provided a detailed analysis

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 921–928, 2008.
© Springer-Verlag Berlin Heidelberg 2008

of four types: (1) univariate analysis, (2) risk index models, (3) multivariate discriminant analysis, and (4) conditional probability model (logit, probit, linear probability models).

Min and Lee [9] were among the first to apply SVM (support vector machines) to bankruptcy prediction problem. Another attempt at hybrid intelligent systems for bankruptcy prediction was made by Tsakonas, Dounias, Doumpos and Zopounidis [10] who developed a model employing neural logic networks through genetic programming. The goal was to obtain the optimal topology of neural networks using genetic programming. A comprehensive survey of research work published between 1968 and 2005 has been compiled by Kumar and Ravi [8]. The paper analyzed a variety of statistical and intelligent methodologies applied to bankruptcy prediction. Organized by technique category, the study concluded that the stand-alone statistical methods were no longer used and neural networks are the most commonly intelligent technique used in the stand-alone mode. However, the author emphasized the potential of hybrid intelligent systems and identified it as the current trend an! d a direction for the future.

The above sampling of the recently published literature on predicting corporate failure shows a vast number of approaches taken to the subject in an attempt to refine the classification model. Most forecasts achieve accuracy between 65% to 85%. Although in the last decade the standard statistical techniques including clustering/classification methods have been largely replaced by more advanced intelligent techniques our study will take a new attempt at subspace clustering algorithm and alter it and adapt for the purpose of accurate prediction.

This paper modifies a fuzzy subspace clustering (FSC) algorithm to a classification algorithm and applies the resulted classification algorithm to bankruptcy prediction. Data clustering provides a number of methods to analyze data sets and extract relevant information from data sets. Technically, data clustering refers to an unsupervised process that divides a given data set into homogeneous groups called clusters such that points within the same cluster are more similar than points across different clusters. An overview of the topic can be found in [5,4].

Unlike data clustering, classification refers to a set of methods that predict the class of an object from attributes or features describing the object [6,7]. In other words, classification relates to the problem of predicting the unknown class of an object. In classification, an object is classified into a pre-defined class using the features that distinguish it from other classes. In general, a classification system consists of two major stages: training and prediction. In the training stage, a training data set is used to train the system; In the prediction stage, the trained system is used to predict the unknown class of an object.

In clustering and classification, a data point or object can be any real item such as a company or a bank branch. An object is characterized by a set of features which are numerical variables such as total assets and total debt. Therefore, an object corresponds to a point in the d-dimensional feature space. A data set is usually described by an $n \times d$ matrix which contains a row for each of the n objects and a column for each of the d features [5,4].

In general, the application of classification for prediction consists of the following steps:

1. Collect data for known objects, i.e. the training set;
2. Train the clustering algorithm using the training set;
3. Use the trained algorithm to classify new objects and to predict their properties.

The remaining part of this paper is organized as follows. In Section 2, we briefly review the FSC method. In Section 3, we present experimental evaluation of the modified FSC algorithm. In Section 4, some conclusions are given.

2 The FSC Method

The FSC algorithm was proposed by Gan et. al. [3] to cluster high-dimensional data sets. In the FSC algorithm, each dimension of the original data is associated with each cluster by a weight. The higher density of a cluster in a dimension, the more weight will be assigned to that dimension.

Before introducing the FSC algorithm, we first define some notations. Let $D = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\} \subset \Re^d$ be a finite data set in the Euclidean space $\Re^d$ and k be an integer $2 \leq k < n$. A fuzzy k-partition of D can be represented by a real $k \times n$ matrix $U = (u_{ji})$ which satisfies

$$u_{ji} \in [0, 1], \quad 1 \leq j \leq k, \, 1 \leq i \leq n, \tag{1a}$$

$$\sum_{j=1}^{k} u_{ji} = 1, \quad 1 \leq i \leq n, \tag{1b}$$

$$\sum_{i=1}^{n} u_{ji} > 0, \quad 1 \leq j \leq k. \tag{1c}$$

A $k \times d$ matrix $W = (w_{jh})$ is said to be a fuzzy dimension weight matrix if W satisfies the following conditions

$$0 \leq w_{jh} \leq 1, \quad 1 \leq j \leq k, \quad 1 \leq h \leq d, \tag{2a}$$

$$\sum_{h=1}^{d} w_{jh} = 1, \quad 1 \leq j \leq k. \tag{2b}$$

The element w_{jh} specifies the probability of the dimension h belonging to the set of cluster dimensions of the cluster j.

Mathematically, the objective function of the FSC algorithm is defined as

$$E_{m,\alpha,\epsilon}(W, Z, U) = \epsilon \sum_{j=1}^{k} \sum_{h=1}^{d} w_{jh}^{\alpha} + \sum_{j=1}^{k} \sum_{i=1}^{n} u_{ji}^{m} \sum_{h=1}^{d} w_{jh}^{\alpha} (x_{ih} - z_{jh})^2, \tag{3}$$

where W, Z and U are the fuzzy dimension weight matrix, the center and the fuzzy k-partition of D, respectively, $m \in (1, \infty)$, $\alpha \in (1, \infty)$ is a weight component or fuzzier, and ϵ is a very small positive real number. It should be noted that any one of W, Z and U can be determined from the other two.

It can be shown that (W^*, Z^*, U^*) is a local minimum of $E_{m,\alpha,\epsilon}$ if and only if, for any $m > 1$, $\alpha > 1$ and $\epsilon > 0$, there holds

$$w_{jh}^* = \frac{1}{\sum_{l=1}^{d} \left[\frac{\sum_{i=1}^{n} (u_{ji}^*)^m (x_{ih} - z_{jh}^*)^2 + \epsilon}{\sum_{i=1}^{n} (u_{ji}^*)^m (x_{il} - z_{jl}^*)^2 + \epsilon} \right]^{\frac{1}{\alpha - 1}}}, \tag{4a}$$

for $1 \le j \le k$ and $1 \le h \le d$,

$$z_{jh}^* = \frac{\sum_{i=1}^{n} (u_{ji}^*)^m x_{ih}}{\sum_{i=1}^{n} (u_{ji}^*)^m}, \tag{4b}$$

for $1 \le j \le k$ and $1 \le h \le d$, and

$$u_{ji}^* = \frac{1}{\sum_{l=1}^{k} \left[\frac{d_{ji}}{d_{li}} \right]^{\frac{1}{m-1}}}, 1 \le j \le k, 1 \le i \le n, \tag{4c}$$

if $d_{li} = \sum_{h=1}^{d} (w_{lh}^*)^\alpha (x_{ih} - z_{lh}^*)^2 > 0$ for all l, i. If $d_{li} = 0$ for some l, i, then u_{ji}^* can be any nonnegative real numbers satisfying

$$\sum_{j=1}^{k} u_{ji}^* = 1, \text{ and } u_{ji}^* = 0 \text{ if } d_{ji} \neq 0. \tag{4d}$$

3 Bankruptcy Prediction by the FSC Method

3.1 Modified FSC

Let $D = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\}$ and $D^* = \{\mathbf{x}_{n+1}, \mathbf{x}_{n+2}, ..., \mathbf{x}_{n+t}\}$ denote the training data set and the test data set, respectively. Objects in the training set are classified into two groups: bankrupted and non-bankrupted. Suppose the first group contains all bankrupted objects and the second group contains all non-bankrupted objects, we then specify the $u_1(\mathbf{x})$ and $u_2(\mathbf{x})$ of the object $\mathbf{x}$ in the training set as

$$u_1(\mathbf{x}) = \begin{cases} 1, \text{ If } \mathbf{x} \text{ is bankrupted;} \\ 0, \text{ Otherwise,} \end{cases} \tag{5}$$

$$u_2(\mathbf{x}) = 1 - u_{1i}. \tag{6}$$

Then we can train the the fuzzy dimension weight matrix and the centers w_{jh} and z_{jh} using the u_{ji}s specified above.

In details, in order to predict whether $\mathbf{y} \in D^*$ is bankrupted, we first need to calculate z_{jh}'s and w_{jh}'s according to Equation (4b) and (4a), respectively. That is,

$$z_{jh} = \frac{\sum\limits_{\mathbf{x}\in D} u_j(\mathbf{x})^m x_h}{\sum\limits_{i=1}^{n} u_j(\mathbf{x})^m}, \tag{7a}$$

for $1 \le j \le 2$ and $1 \le h \le d$, where $u_j(\mathbf{x})$'s are defined in Equation (5) and Equation (6), and

$$w_{jh} = \frac{1}{\sum\limits_{l=1}^{d} \left[\frac{\sum\limits_{\mathbf{x}\in D} u_j(\mathbf{x})^m (x_h - z_{jh})^2 + \epsilon}{\sum\limits_{\mathbf{x}\in D} u_j(\mathbf{x})^m (x_l - z_{jl})^2 + \epsilon} \right]^{\frac{1}{\alpha-1}}}, \tag{7b}$$

for $1 \le j \le 2$ and $1 \le h \le d$, where u_{ji}'s are defined in Equation (5) and Equation (6), and z_{jh}'s are calculated as in Equation (7a). Using the z_{jh}'s in Equation (7a) and w_{jh}'s in Equation (7b), we calculate the memberships of $\mathbf{y}$ as

$$\begin{aligned} u_1(\mathbf{y}) &= \frac{1}{1 + \left[\frac{d_1(\mathbf{y})}{d_2(\mathbf{y})} \right]^{\frac{1}{m-1}}}, \\ u_2(\mathbf{y}) &= 1 - u_1(\mathbf{y}), \end{aligned} \tag{8}$$

where

$$d_l(\mathbf{y}) = \sum_{h=1}^{d} w_{lh}^{\alpha} (y_h - z_{lh})^2, \quad l = 1, 2,$$

with w_{lh}'s and z_{lh}'s being defined in Equation (7b) and Equation (7a), respectively. Here we assume that $d_1(\mathbf{y}) > 0$ and $d_2(\mathbf{y}) > 0$. For real data set, this assumption is reasonable. Based on $u_1(\mathbf{y})$ calculated above, $\mathbf{y}$ is classified as follows. If $u_1(\mathbf{y}) > 0.5$ (equivalently $d_1(\mathbf{y}) < d_2(\mathbf{y})$), $\mathbf{y}$ is classified into the bankrupted group; if $u_1(\mathbf{y}) < 0.5$ (equivalently $d_1(\mathbf{y}) > d_2(\mathbf{y})$), $\mathbf{y}$ is classified as the non-bankrupted group; otherwise, the classification is inconclusive.

Once we classified an object in D^*, we move this object from the test data set to the training data set D and and update w_{jh}'s and z_{jh} according to Equation (7b) and Equation (7a), respectively. Then we repeat the aforementioned process to classify another object in D^* until all objects in the test data set D^* have been classified. The pseudo code of the algorithm is described in Algorithm 1.

3.2 The Data Set

The data set includes 303 companies and 28 of them went bankrupted one year later. Each company is described by 10 attributes, which includes total assets (TA), working capital (WC), earnings before income, tax, depreciation

Algorithm 1. The pseudo code of modified FSC for prediction of bankruptcy

Require: D – the training data set, D^* – the test data set;
repeat
  Calculate z_{1h}'s and z_{2h}'s according to Equation (7a);
  Calculate w_{1h}'s and w_{2h}'s according to Equation (7b);
  Calculate $u_1(\mathbf{y})$ according to Equation (8);
  Classify $\mathbf{y}$;
  if $u_1(\mathbf{y}) \neq 0.5$ **then**
    Move $\mathbf{y}$ from D^* to D;
  else
    Remove $\mathbf{y}$ from D^*;
  end if
until D^* is empty

and amortization (EBITDA), retained earnings (RE), shareholders equity (EQ), total current liabilities (CL), interest expense (IN), cash flow from operations (CF), stability of earnings (SE) and total liabilities (TL).

3.3 Experimental Evaluation

The modified FSC algorithm is coded in the MATLAB script language. Our experiments are conducted on a PC with 1.7G CPU and 512M RAM. In our experiments, we specify $\alpha = 2$, $m = 2$ and $\epsilon = 0.0001$.

For the given data set, we randomly select part of the data as training data and the remaining data as test data. Since the training data is selected randomly from the whole data set, we run the algorithm 100 times and calculate the average accuracy. The average accuracy is calculated as follows. Let p be the percentage of training data and n the number of objects in the training data set, then we have $n = [275p] + [28p]$, where $[a]$ denotes the largest integer less than or equal to a. Therefore, the number of objects in the test data set is $n_A = 303 - n$. Let $n_B = 28 - [28p]$ denote the number of bankrupted objects in the test data set, n_C be the number of objects classified correctly, and n_D the number of bankrupted objects classified correctly. The total accuracy and accuracy (Bankrupted) are defined as

$$R = \frac{n_C}{n_A},$$

and

$$r = \frac{n_D}{n_B},$$

respectively. The average total accuracy and average accuracy (bankrupted) of 100 runs are defined as

$$\bar{R} = \sum_{i=1}^{100} \frac{R_i}{100},$$

and

$$\bar{r} = \sum_{i=1}^{100} \frac{r_i}{100},$$

Table 1. Average prediction accuracy of 100 runs of the modified FSC algorithm with various percentages of training data

Training Data	Avg. Accuracy (Total)	Avg. Accuracy (Bankrupted)	Avg. Accuracy (Non-bankrupted)
95%	91.19%	35.00%	99.21%
90%	93.58%	40.33%	99.29%
85%	92.40%	38.20%	98.86%
80%	92.66%	34.33%	99.02%
75%	93.00%	34.71%	98.91%
70%	92.79%	34.11%	99.16%
50%	92.95%	33.36%	98.99%
20%	89.20%	26.09%	95.80%
10%	78.75%	21.62%	84.74%

respectively, where R_i and r_i are total accuracy and accuracy (Bankrupted) for the ith run.

The classification results of the algorithm are given in Table 1. We can easily observe that the proposed method provides an impressively high prediction accuracy for the non-bankrupted companies. However, the prediction accuracy rate for the bankrupted companies is farily low. This is not a surprising result since the data set only includes less than 10% bankrupted companies and the training process can not accurately recognize the features of the bankruptcy. If a balanced data set is used, the proposed algorithm will produce a very high prediction rate for both bankrupted and non-bankrupted companies. Nevertheless, the current result still provides significant insights to the industry. Our models' high levels of non-bankruptcy prediction accuracy can translate into large financial and other benefits to organizations through such activities as credit approval, and loan portfolio and security management.

4 Conclusion

In this paper we presented a fuzzy classification algorithm and used it for the purpose of bankruptcy prediction. The presented algorithm was modified from a subspace-clustering algorithm and adapted for object categorization. The most important advantage of our algorithm is that this it is able to classify an object without the need for feature selection. The implementation (empirical testing) of our algorithm demonstrates an exceptional and consistent predictability ratio of non-bankrupted objects averaging at about 99%. In terms of total accuracy, reaching 93%, the results produced by our algorithm are above average. Due to the small percentage of bankrupted companies (28 bankrupted objects among 303 objects), the results produced by our algorithm are not high in accuracy for bankrupted objects. We claim that given a more balanced data set with an appropriate level of bankrupted objects, an empirical test performed on this new data using modified FSC would yield an eq! uivalently high performance.

Our FSC algorithm affords opportunities for future research. The proposed methodology can be extended and used with other classification algorithm as a hybrid system to amplify the advantages of the algorithm and further improve its classification performance.

References

1. Balcean, S., Ooghe, H.: 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. The British Accounting Review 38(1), 63–93 (2006)
2. Beaver, W.: Financial ratios predictors of failure. Empirical research in accounting: selected studies 1966. Journal of Accounting Research (Suppl. 4), 71–111 (1967)
3. Gan, G., Wu, J., Yang, Z.: A fuzzy subspace algorithm for clustering high dimensional data. In: Li, X., Zaïane, O.R., Li, Z. (eds.) ADMA 2006. LNCS (LNAI), vol. 4093, pp. 271–278. Springer, Heidelberg (2006)
4. Gordon, A.: A review of hierarchical classification. Journal of the Royal Statistical Society. Series A (General) 150(2), 119–137 (1987)
5. Jain, A., Murty, M., Flynn, P.: Data clustering: A review. ACM Computing Surveys 31(3), 264–323 (1999)
6. Jain, A., Duin, R.P.W., Mao, J.: Statistical pattern recognition: a review. IEEE Transactions on Pattern Analysis and Machine Intelligence 322(1), 4–37 (2000)
7. Kulkarni, S.R., Lugosi, G., Venkatesh, S.S.: Learning pattern classification-a survey. IEEE Transactions on Information Theory 44(6), 2178–2206 (1998)
8. Ravi Kumar, P., Ravi, V.: Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. European Journal of Operational Research 180(1), 1–28 (2007)
9. Min, J.H., Lee, Y.C.: Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert Systems With Applications 28(4), 603–614 (2005)
10. Tsakonas, A., Dounias, G., Doumpos, M., Zopounidis, C.: Bankruptcy prediction with neural logic networks by means of grammar-guided genetic programming. Expert Systems With Applications 30(3), 449–461 (2006)

Financial Time Series Analysis of SV Model by Hybrid Monte Carlo

Tetsuya Takaishi

Hiroshima University of Economics,
Hiroshima 731-0192 Japan
takaishi@hiroshima-u.ac.jp

Abstract. We apply the hybrid Monte Carlo (HMC) algorithm to the financial time sires analysis of the stochastic volatility (SV) model for the first time. The HMC algorithm is used for the Markov chain Monte Carlo (MCMC) update of volatility variables of the SV model in the Bayesian inference. We compute parameters of the SV model from the artificial financial data and compare the results from the HMC algorithm with those from the Metropolis algorithm. We find that the HMC decorrelates the volatility variables faster than the Metropolis algorithm. We also make an empirical analysis based on the Yen/Dollar exchange rates.

Keywords: Hybrid Monte Carlo Algorithm, Stochastic Volatility Model, Markov Chain Monte Carlo, Bayesian Inference, Financial Data Analysis.

1 Introduction

It is well known that financial time series of asset returns shows various interesting properties which can not be explained from the assumption of that the time series obeys the Brownian motion. Those properties are classified as stylized facts[1]. Some examples of the stylized facts are (i) fat-tailed distribution of return (ii) volatility clustering (iii) slow decay of the autocorrelation time of the absolute returns. The true dynamics behind the stylized facts is not fully understood. There are some attempts to construct physical models based on spin dynamics[2]-[5] and they are able to capture some of the stylized facts.

The volatility is an important value to measure the risk in finance. The volatilities of asset returns change in time and shows the volatility clustering. In order to mimic these empirical properties of the volatility Engle advocated the autoregressive conditional hetroskedasticity (ARCH) model[6] where the volatility variable changes deterministically depending on the past squared value of the return. Later the ARCH model is generalized by adding also the past volatility dependence to the volatility change. This model is known as the generalized ARCH (GARCH) model[7]. The parameters of the GARCH model applied to financial time series are easily determined by the maximum likelihood method.

The stochastic volatility (SV) model[8,9] is another model which captures the properties of the volatility. Contrast to the GARCH model of which the volatility change is deterministic, the volatility of the SV model changes stochastically in

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 929–936, 2008.
© Springer-Verlag Berlin Heidelberg 2008

time. As a result the likelihood function of the SV model is given as a multiple integral of the volatility variables. Such an integral in general is not analytically calculable and thus the determination of the parameters in the SV model by the maximum likelihood method becomes ineffective.

For the SV model the MCMC method based on the Bayesian approach is developed. In the MCMC of the SV model one has to update the parameter variables and also the volatility ones. The number of the volatility variables to be updated increases with the data size of the time series. Usually the update scheme of the volatility variables is based on the local one such as the Metropolis-type algorithm[8]. It is however known that when the local update scheme is done for the volatility variables which have interactions to their neighbor variables in time, the autocorrelation time of sampled volatility variables becomes high and thus the local update scheme is not effective. In order to improve the efficiency of the local update method the blocked scheme which updates several variables at once is also proposed[10].

In this paper we use the HMC algorithm[11] to update the volatility variables. There exists an application of the HMC algorithm to the GARCH model[12] where three GARCH parameters are updated by the HMC scheme. It is more interesting to apply the HMC for update of the volatility variables because the HMC algorithm is a global update scheme which can update all variables at once. To examine the HMC we calculate the autocorrelation function of the volatility variables and compare the result with that of the Metropolis algorithm.

2 Stochastic Volatility Model and Its Bayesian Inference

2.1 Stochastic Volatility Model

The standard version of the SV model[8,9] is given by

$$y_t = \sigma_t \epsilon_t = \exp(h_t/2)\epsilon_t, \tag{1}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \tag{2}$$

where $y_t = (y_1, y_2, ..., y_n)$ represents the time series data, h_t is defined by $h_t = \ln \sigma_t^2$ and σ_t is called volatility. The error terms ϵ_t and η_t are independent normal distributions $N(0,1)$ and $N(0,\sigma_\eta^2)$ respectively.

For this model the parameters to be determined are μ, ϕ and σ_η^2. Let us use θ as $\theta = (\mu, \phi, \sigma_\eta^2)$. The likelihood function $L(\theta)$ for the SV model is written as

$$L(\theta) = \int \prod_{t=1}^{n} f(\epsilon_t|\sigma_t^2) f(h_t|\theta) dh_1 dh_2 ... dh_n, \tag{3}$$

where

$$f(\epsilon_t|\sigma_t^2) = (2\pi\sigma_t^2)^{-\frac{1}{2}} \exp(-\frac{y_t^2}{2\sigma_t^2}), \tag{4}$$

$$f(h_1|\theta) = (\frac{2\pi\sigma_\eta^2}{1-\phi^2})^{-\frac{1}{2}} \exp(-\frac{[h_1-\mu]^2}{2\sigma_\eta^2/(1-\phi^2)}), \tag{5}$$

$$f(h_t|\theta) = (2\pi\sigma_\eta^2)^{-\frac{1}{2}} \exp(-\frac{[h_t - \mu - \phi(h_{t-1} - \mu)]^2}{2\sigma_\eta^2}). \quad (6)$$

2.2 Bayesian Inference for the SV Model

In the Bayesian theorem, the probability distributions of the parameters to be estimated are given by

$$f(\theta|y) = \frac{1}{Z} L(\theta)\pi(\theta), \quad (7)$$

where Z is the normalization constant $Z = \int L(\theta)\pi(\theta)d\theta$ and $\pi(\theta)$ is a prior distibution of θ for which we make a certian assumption. The values of the parameters are inferred as the expectation values of θ given by $\langle\theta\rangle = \int \theta f(\theta|y)d\theta$. In general this integral can not be performed analytically. For that case, one can use the MCMC method to estimate the expectation values numerically. In the MCMC method, we first generate a series of θ with a probability $P(\theta) = f(\theta|y)$. Let $\theta^{(i)} = (\theta^{(1)}, \theta^{(2)}, ..., \theta^{(k)})$ be values of θ generated by a MCMC sampling. Then using these values the expectation value of θ is estimated by $\langle\theta\rangle = \frac{1}{k}\sum_{i=1}^{k} \theta^{(i)}$.

For the SV model, in addition to θ, volatility variables h_t also have to be updated since they are integrated out as in eq.(3). Let $P(\theta, h_t)$ be the joint probability distribution of θ and h_t. Then $P(\theta, h_t)$ is given by

$$P(\theta, h_t) \sim \bar{L}(\theta, h_t)\pi(\theta), \quad (8)$$

where $\bar{L}(\theta, h_t) = \prod_{t=1}^{n} f(\epsilon_t|h_t)f(h_t|\theta)$.

For the prior $\pi(\theta)$ we assume that $\pi(\sigma_\eta^2) \sim (\sigma_\eta^2)^{-1}$ and for others $\pi(\mu) = \pi(\phi) = constant$. The probability distributions for the parameters and the volatility variables are derived from eq.(8)[8,9]. The probability distributions and their update schemes are given in the followings.

- σ_η^2 update scheme.
  The probability distribution of σ_η^2 is given by

$$P(\sigma_\eta^2) \sim (\sigma_\eta^2)^{-\frac{n}{2}-1} \exp\left(-\frac{A}{\sigma_\eta^2}\right), \quad (9)$$

  where $A = \frac{1}{2}\{(1-\phi^2)(h_1 - \mu)^2 + \sum_{t=2}^{n}[h_t - \mu - \phi(h_{t-1} - \mu)]^2\}$.
  Since eq.(9) is an inverse gamma distribution we can easily draw a value depending on eq.(9) by using an appropriate statistical library.
- μ update scheme.
  The probability distribution of μ is given by

$$P(\mu) \sim \exp\left\{-\frac{B}{2\sigma_\eta^2}(\mu - \frac{C}{B})^2\right\}, \quad (10)$$

  where $B = (1-\phi^2) + (n-1)(1-\phi)^2$,
  and $C = (1-\phi^2)h_1 + (1-\phi)\sum_{t=2}^{n}(h_t - \phi h_{t-1})$.
  Eq.(10) is a Gaussian distribution. Again we can easily update μ.

- ϕ update scheme.
  The probability distribution of ϕ is given by

$$P(\phi) \sim (1-\phi^2)^{1/2} \exp\{-\frac{D}{2\sigma_\eta^2}(\phi - \frac{E}{D})^2\}, \tag{11}$$

  where $D = -(h_1-\mu)^2 + \sum_{t=2}^{n}(h_{t-1}-\mu)^2$ and $E = \sum_{t=1}^{n}(h_t-\mu)(h_{t-1}-\mu)$. In order to update ϕ with eq.(11), we use the Metropolis-Hastings algorithm [13,14]. Let us write eq.(11) as $P(\phi) \sim P_1(\phi)P_2(\phi)$ where

$$P_1(\phi) = (1-\phi^2)^{1/2}, \tag{12}$$

$$P_2(\phi) \sim \exp\{-\frac{D}{2\sigma_\eta^2}(\phi - \frac{E}{D})^2\}. \tag{13}$$

  Since $P_2(\phi)$ is a Gaussian distribution we can easily draw ϕ from eq.(13). Let ϕ_{new} be a candidate given from eq.(13). Then in order to obtain the correct distribution, ϕ_{new} is accepted with the following probability P_{MH}.

$$P_{MH} = \min\left\{\frac{P(\phi_{new})P_2(\phi)}{P(\phi)P_2(\phi_{new})}, 1\right\} = \min\left\{\sqrt{\frac{(1-\phi_{new}^2)}{(1-\phi^2)}}, 1\right\}. \tag{14}$$

  In addition to the above step we restrict ϕ within $[-1,1]$ to avoid a negative value in the calculation of square root.
- Probability distribution for $P(h_t)$.
  The probability distribution of the volatility variables h_t is given by

$$P(h_1, h_2, ..., h_n) \sim \tag{15}$$
$$\exp\left(-\sum_{i=1}^{n}\{\frac{h_t}{2} + \frac{\epsilon_t^2}{2}e^{-h_t}\}\frac{[h_1-\mu]^2}{2\sigma_\eta^2/(1-\phi^2)} - \sum_{i=2}^{n}\frac{[h_t-\mu-\phi(h_{t-1}-\mu)]^2}{2\sigma_\eta^2}\right).$$

  This probability distribution is not a simple function for drawing values of the volatility variables h_t. A conventional method is the Metropolis method[13] which updates the variables locally. Here we use the HMC algorithm which updates the volatility variables globally.

3 Hybrid Monte Carlo Algorithm

The HMC algorithm is originally developed for the MCMC of the lattice Quantum Chromo Dynamics (QCD) calculations [11] where local type update algorithms are not effective. The notable feature of the HMC algorithm is that it updates a number of variables simultaneously.

Here we briefly describe the HMC algorithm. The HMC algorithm combines molecular dynamics (MD) simulations and the Metropolis test. Let $f(x)$ be a probability density and $O(x)$ a function of $x = (x_1, x_2, ..., x_n)$. We determine the expectation value of $O(x)$ with the probability density $f(x)$ which is given by

$$\langle O(x)\rangle = \int O(x)f(x)dx = \int O(x)e^{lnf(x)}dx. \tag{16}$$

Now let us introduce momentum variables $p = (p_1, p_2, ..., p_n)$ conjugate to the variables x and rewrite eq.(16) as

$$\langle O(x) \rangle = \frac{1}{Z} \int O(x) e^{-\frac{1}{2}p^2 + lnf(x)} dxdp = \frac{1}{Z} \int O(x) e^{-H(p,x)} dxdp. \tag{17}$$

where $Z = \int e^{-\frac{1}{2}p^2} dp$. $H(p, x)$ is the Hamiltonian defined by $H(p, x) = \frac{1}{2}p^2 - lnf(x)$ where p^2 stands for $\sum_{i=1}^{n} p_i^2$. The introduction of p does not change the value of $\langle O(x) \rangle$.

In the HMC algorithm, new candidates of the variables are drawn by integrating the Hamilton's equations of motions. The Hamilton's equations of motions are solved numerically by doing the MD simulation with a fictitious time. To solve the equations we use the standard 2nd order leapfrog integrator. One could use improved integrators[15] or higher order integrators[16,17] if necessary.

Let (p', x') be the new candidates given by the MD simulation. The new candidates are accepted with a probability $\min\{1, \exp(-\Delta H)\}$ where $\Delta H = H(p', x') - H(p, x)$. Since the Hamilton's equation of motions are not solved exactly ΔH deviates from zero. The magnitude of the deviation is tuned by the discrete time step size in the MD simulation such that the acceptance of the new candidates becomes high.

For the volatility variables of the SV model, from eq.(15) the Hamiltonian can be defined by

$$H(p_t, h_t) = \sum_{i=1}^{n} \frac{1}{2}p_i^2 + \sum_{i=1}^{n} \{\frac{h_i}{2} + \frac{\epsilon_i^2}{2} e^{-h_i}\} + \frac{[h_1 - \mu]^2}{2\sigma_\eta^2/(1-\phi^2)} + \sum_{i=2}^{n} \frac{[h_i - \mu - \phi(h_{i-1} - \mu)]^2}{2\sigma_\eta^2}, \tag{18}$$

where p_i is defined as a conjugate momentum to h_i.

4 Numerical Test

In this section we investigate the HMC algorithm for the SV model with artificial financial data. The artificial data is generated with a set of known parameters. We try to infer the values of those parameters by the HMC and Metropolis algorithms and compares the results.

Using eq.(1) with $\phi = 0.97$,$\sigma_\eta^2 = 0.05$ and $\mu = -1$ we have generated 2000 data. To this data we made the Bayesian inference with the HMC and Metropolis algorithms. The initial parameters are set to $\phi = 0.5$,$\sigma_\eta^2 = 1.0$ and $\mu = 0$. The first 10000 samples are discarded as thermalization or burn-in process. Then 200000 samples are recorded for analysis. The acceptance of the volatility variables is tuned to be about 50%.

Fig.1 shows the history of the volatility variable h_{100}. We use h_{100} as the representative one of the volatility variable. We observe the similar behavior for other volatility variables. As seen in Fig.1 the correlation of the volatility variable from the HMC algorithm is smaller than that from the Metropolis algorithm. To quantify this we calculated the autocorrelation function (ACF) of the volatility

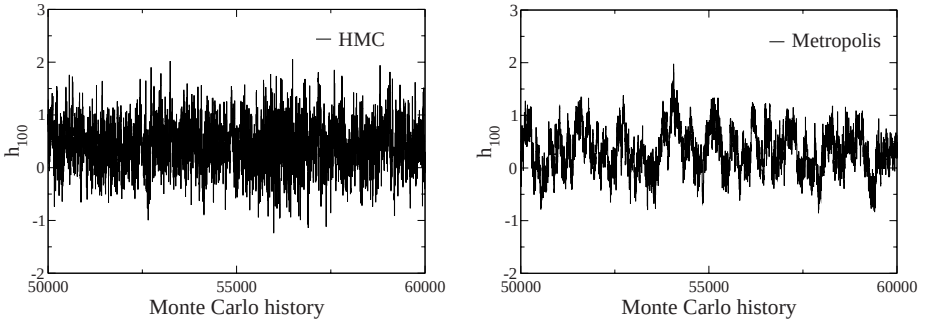


Fig. 1. Monte Carlo history of HMC (left) and Metropolis (right) in the window from 50000 to 60000 Monte Carlo history

variable shown in Fig.2. The ACF is defined as

$$ACF(t) = \frac{\frac{1}{N}\sum_{j=1}^{N}(x(j)- < x >)(x(j+t)- < x >)}{\sigma_x^2}, \tag{19}$$

where $< x >$ and σ_x^2 are the average value and the variance of x respectively.

The autocorrelation time τ_{int} of the volatility variables is given in Table 1. The values in the parentheses represent the errors estimated by the jackknife method. The autocorrelation time is defined by $\tau_{int} = \frac{1}{2} + \sum_{t=1}^{\infty} ACF(t)$.

The HMC algorithm gives a smaller autocorrelation time than the Metropolis algorithm, which means that the HMC algorithm samples the volatility variables more effectively than the Metropolis algorithm.

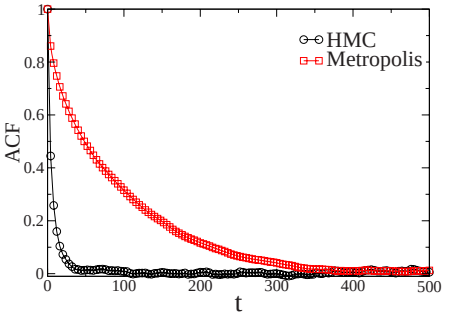


Fig. 2. Autocorrelation function of the volatility variable h_{100} for the HMC and Metropolis algorithms. We see that the ACF from HMC decreases quickly as the Monte Carlo time increases.

The autocorrelation times for the parameters of the SV model are summarized in Table 1. The autocorrelation times from the HMC algorithm are similar to those of the Metropolis algorithm except for τ_{int} of μ.

Table 1. Results estimated by the HMC and Metropolis algorithms

	ϕ	μ	σ_η^2	h_{100}
true	0.97	-1	0.05	
HMC	0.978(7)	-0.92(26)	0.053(12)	
τ_{int}	540(60)	3(1)	1200(150)	18(1)
Metropolis	0.978(7)	-0.92(26)	0.052(11)	
τ_{int}	400(100)	13(2)	1000(270)	210(50)

Table 2. Results estimated by the HMC to the Yen/Dollar exchange rates

	ϕ	μ	σ_η^2	h_{100}
HMC	0.960(12)	-1.13(8)	0.014(4)	
τ_{int}	610(300)	14(6)	1400(800)	55(11)

The values of the SV parameters estimated by the HMC and the Metropolis algorithms are given in Table 1. The values in the parentheses represent the standard deviations of the sampled data. The results from the both algorithms well reproduce the true values used for the generation of the artificial financial data. Furthermore for each parameter two values obtained by the HMC and the Metropolis algorithms agree well. This is not surprising because the same data is used for the both calculations by the HMC and Metropolis algorithms.

5 Empirical Study

We have also made an empirical study of the SV model by the HMC. The empirical study is based on daily data of the exchange rates for Japanese Yen and US dollar. The sampling period is 1 March 2000 to 29 February 2008, which has 2007 observations. The exchange rates p_i are transformed to $r_i = 100\ln(p_i/p_{i-1} - \bar{s})$ where $\bar{s}$ is the average value of $\ln(p_i/p_{i-1})$. The MCMC sampling is performed as in the previous section. The first 10000 MC samples are discarded and then 20000 samples are recoded for the analysis. The estimated values of the parameters are summarized in Table 2. The estimated value of ϕ is close to one, which means the persistency of the volatility shock. The similar values are obtained in the previous studies[8,9].

6 Summary

The HMC algorithm is applied for the Bayesian inference of the SV model. It is found that the correlations of the volatility variables sampled by the HMC algorithm are much reduced. On the other hand we observe no significant improvement on the correlations of the sampled parameters of the SV model. Thus it is concluded that the HMC algorithm has a similar efficiency to the Metropolis algorithm and it is an alternative algorithm for the Bayesian inference of the SV model.

If one needs to calculate a certain quantity depending on the volatility variables, then the HMC algorithm may serve as a good algorithm which samples the volatility variables effectively because the HMC algorithm decorrelates the sampled volatility variables faster than the Metropolis algorithm.

Acknowledgments

The numerical calculations were carried out on SX8 at the Yukawa Institute for Theoretical Physics in Kyoto University and on Altix at the Institute of Statistical Mathematics. The author thanks K.Maekawa, Y.Tokutsu, T.Morimoto, K.Kawai, K.Tei and X.H. Lu for valuable discussions.

References

1. Cont, R.: Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. Quantitative Finance 1, 223–236 (2001)
2. Bornholdt, S.: Expectation Bubbles in a Spin Model of Markets: Intermittency from Frustration across Scales. Int. J. Mod. Phys. C 12, 667–674 (2001)
3. Sznajd-Weron, K., Weron, R.: A Simple Model of Price Formation. Int. J. Mod. Phys. C 13, 115–123 (2002)
4. Kaizoji, T., et al.: Dynamics of Price and Trading Volume in a Spin Model of Stock Markets with Heterogeneous Agents. Physica A 316, 441–452 (2002)
5. Takaishi, T.: Simulations of Financial Markets in a Potts-like Model. Int. J. Mod. Phys. C 13, 1311–1317 (2005)
6. Engle, R.F.: Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom inflation. Econometrica 60, 987–1007 (1982)
7. Bollerslev, T.: Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, 307–327 (1986)
8. Jacquier, E., Polson, N.G., Rossi, P.E.: Bayesian Analysis of Stochastic Volatility Models. Journal of Business & Economic Statistics 12, 371 (1994)
9. Kim, S., Shephard, N., Chib, S.: Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. Review of Economic Studies 65, 361–393 (1998)
10. Shephard, N., Pitt, M.K.: Likelihood Analysis of Non-Gaussian Measurement Time Series. Biometrika 84, 653–667 (1994)
11. Duane, S., et al.: Hybrid Monte Carlo. Phys. Lett. B 195, 216–222 (1987)
12. Takaishi, T.: Bayesian Estimation of GARCH model by Hybrid Monte Carlo. In: Proceedings of the 9th Joint Conference on Information Sciences 2006, CIEF-214, p. 159 (2006), doi:10.2991/jcis
13. Metropolis, N., et al.: Equations of State Calculations by Fast Computing Machines. J. of Chem. Phys. 21, 1087–1091 (1953)
14. Hastings, W.K.: Monte Carlo Sampling Methods Using Markov Chains and Their Applications. Biometrika 57, 97–109 (1970)
15. Takaishi, T., de Forcrand, P.: Testing and Tuning Symplectic Integrators for Hybrid Monte Carlo Algorithm in Lattice QCD. Phys. Rev. E 73, 036706 (2006)
16. Takaishi, T.: Choice of Integrators in the Hybrid Monte Carlo Algorithm. Comput. Phys. Commun. 133, 6–17 (2000)
17. Takaishi, T.: Higher Order Hybrid Monte Carlo at Finite Temperature. Phys. Lett. B 540, 159–165 (2002)

The Predicted Model of International Roughness Index for Drainage Asphalt Pavement

Chien-Ta Chen[1], Ching-Tsung Hung[2], Chien-Cheng Chou[1], Ziping Chiang[3], and Jyh-Dong Lin[1]

[1] National Central University, Department of Civil Engineering, No.300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan (R.O.C.)
[2] Kainan University, Department of Transportation Technology and Supply Chain, No. 1 Kainan Road , Luchu, Taoyuan County 33857, Taiwan (R.O.C.)
[3] Leader University, Department of Logistics Management, No.188, Sec. 5, Anjhong Rd., Tainan City 709, Taiwan (R.O.C.)
{953402003, ccchou, jyhdongl}@cc.ncu.edu.tw, cthung@mail.knu.eud.tw, ziping@mail.leader.edu.tw

Abstract. Roughness is a very important performance indicator in pavement maintenance management systems. The international rough index (IRI) of indicators was used as the core of roughness measuring of degree gradually in recent years domestically, but to predict a smooth degree of indicators the method is set up differently as there is a different prediction indicator because the theoretical foundation is different. This research bases its measurement data on the porous asphalt pavement of national highway number 3. We predicts the deterioration of IRI by different ways, including grey forecast, multiple regression, genetic programming. The result of this research found that there are better results in genetic programming method prediction than roughness index foundation method. The reason that it accords with the best prediction result is that heredity can carry on the change to plan in the parameter.

Keywords: International roughness index, grey forecast, pluralism regression, genetic programming, and multiple regressions.

1 Introduction

The roughness of road surface is an important indicator in determining road rideability. The unevenness of the road surface could increase the vertical stress on the road and accelerate the development of fatigue damage. Hence, the unevenness of the road surface, without doubt, could speed up the deterioration of the pavement.

The prediction modes of pavement roughness can be categorized into two models: determinism model and probability theory model. The former is the regression model, such as the prediction model brought up by Mrawira and Wile [1] which uses historical data as the basis to predict road performance and explains the relation between the interaction of traffic load and climate and the structural strength of the pavement and the roughness. The probability theory model includes a random process of Markov chain and neural network model. Yu Cheng-hsin [2] uses a Markov chain as the basic modeling structure and adopts concepts of the queuing theory to simulate and create a

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 937 – 945, 2008.
© Springer-Verlag Berlin Heidelberg 2008

dynamic analysis model for pavement performance. Hung Ching-tsung [3] uses the neural network to create a serviceability prediction model.

Considering the problems of past prediction modes, the research uses the grey forecast model and genetic programming model to create a roughness prediction mode and compares with the traditional multiple regression model. Section two gives an introduction to gray forecast principles and genetic programming methods. Section three formulates how to design and survey a drainage asphalt pavement. Section four discusses the predictability of different prediction methods. The last section is the conclusion of this research.

2 Prediction Methods

2.1 Gray Forecast Model

The Grey forecast (Grey Model, GM) was proposed in 1982 by Deng Ju-long [4-7] to analyze the properties of a system when the system model is unclear and information is incomplete. Grey model has been applied in different fields of research over the years. Grey model provides thinking and solutions to problems with information partly known. The more data collected, the easier it would be to detect the statistic properties in the random processing. In the grey model theory, the interference of the environment will disorder discrete function data, which is the eigenvalue of the system behavior. But since the system is complete, the disordered explicit data must contain some internal regularity. Using grey processing to sort raw data can weaken the randomness and strengthen regularity, and create a model for prediction or forecast.

Grey forecast uses the sequence produced by the system, and computes through accumulated generation operation. Using pseudo differential equation, i.e., grey model, for prediction often adopts Grey Model (1,1). This model is supposed that the sequence of original data is expressed as X={x(t)| t=1, 2, ..., n}. After making the first order accumulation, we obtain the generation model block: Z={z(t)| t=1, 2, ..., n},where

$$z_{(t)} = \sum_{k=1}^{t} x_{(t)} . \tag{1}$$

Because of the original model of GM(1,1) cannot carry out successive analysis and prediction for a time series. In another words, we cannot use it as a truly difference equation. Thus we use the following difference equation instead of the original model of GM(1,1).

$$\frac{dz_{(t)}}{dt} + az_{(t)} = b \tag{2}$$

where "a" and „b‰are determined by the given data(Deng (1989). We can find the coefficients "a" and „b‰by the follows.

$$\theta = \begin{bmatrix} a \\ b \end{bmatrix} = [Q^T Q]^{-1} Q^T P \tag{3}$$

where, $P = \begin{bmatrix} x(2) \\ x(3) \\ \vdots \\ x(n) \end{bmatrix}$, $Q = \begin{bmatrix} -\frac{1}{2}(z_{(1)} + Z_{(2)}) & 1 \\ \cdots & \\ -\frac{1}{2}(z_{(n-1)} + Z_{(n)}) & 1 \end{bmatrix}$

Solving the differential Eq. (2) gives the prediction model of the generation model block Z.

$$\hat{z}_{(t)} = [x_{(1)} - \frac{b}{a}]e^{-a(t-1)} + \frac{b}{a}, (t=1,2,\ldots,n) \tag{4}$$

And we will know the accumulated subtraction returning to original sequence X, which gives the prediction model.

$$\begin{aligned} \hat{x}_{(1)} &= x_{(1)} \\ \hat{x}_{(t)} &= [x_{(1)} - \frac{b}{a}](1 - e^{a})e^{-a(t-1)}, (t=2,3,\ldots,n) \end{aligned} \tag{5}$$

2.2 Genetic Programming

Genetic programming is an algorithm based methodology inspired by biological evolution to find an automated production and computer program that performs a user-defined task. Genetic programming is a special machine learning technology that uses evolutionary algorithms. It begins with a group of randomly produced individuals comprising computer programs and uses the ability to complete a task defined in a process to determine the goodness of fit of a certain process. It employs Darwin's natural selection (survival of the fittest) to determine the process.

Genetic programming was first proposed by Smith and Cramer [8]. In 1992, Koza [9] published Genetic Programming: An Introduction: On the Automatic Evolution of Computer Programs and Its Applications as an introduction to genetic programming.

When genetic programming is used to produce code, it is necessary to define the function set and terminal set that will be used in the program tree and the function in the set should be computed with suitable parameters. In the program tree, the function nodes are all internal nodes and the parameters in the computation process are the results of the child tree. The function set can include the following functions: calculation operators (+, -, *etc.), mathematic function（sin, cos, exp, log, etc.）, Boolean

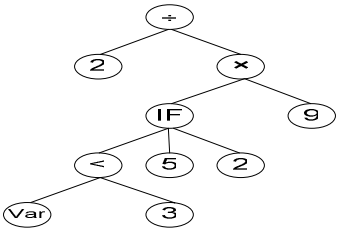


Fig. 1. Tree structure of genetic programming

operators（and, or, not, etc）, if-then-else, do-until, recursive function, and functions used in other professional fields.

The terminal tends to be a variable or constant, or a function that requires no parameter in the computing process. In the program tree, such nodes are the terminal nodes, or the leaves. Besides, it is necessary to define a function that is used to evaluate the goodness of fit of each program. Like genetic algorithms, genetic programming has three operators that can produce children: crossover, reproduction, and mutation. The following is the implementation process of the three operators:

Crossover: in the computation process, two program individuals choose a branch through random number generation and are placed into the tree. The chosen branches then exchange to produce a new generation (Figure 2).

Reproduction: in this process, a program is reproduced to a next generation and the reproduced parent and its child's program are completely identical.

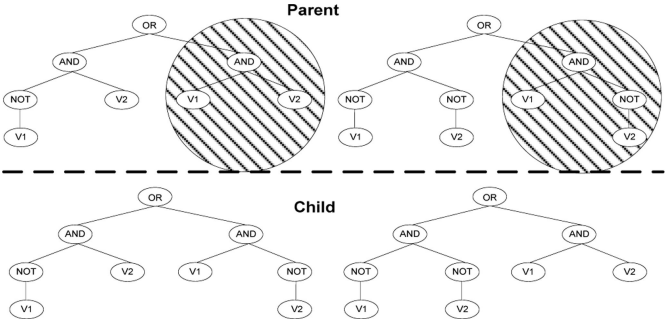


Fig. 2. Crossover (the upper half is the development "before the crossover"; the grey section is the child tree randomly chosen; and the lower half is the development "after the crossover")

Mutation: in this process, a branch of a program individual is replaced by a new branch randomly generated.

2.3 Multiple Regression

Multiple regression analysis, also known as pluralism regression, is one of the methods commonly used for quantitative prediction. It uses the cause-and-effect relation of the changes of the internal factors to predict an event's future development. Since it is based on the regularity of an event's internal development, it can produce more precise results.

Multiple linear regression means the linear relationship between a dependent variable and many independent variables. When using multiple regressions, it is necessary to use univariate analysis and multivariate statistical analysis to make sure each of the variables to be input for analyzing conforms to the basic assumptions of OLS linear regression analysis. After the regression mode is chosen, the parameter prediction and goodness of fit should be evaluated. Before we seriously consider the analysis results, we should diagnose the residuals. We often diagnose

the residuals after we make sure the specification of the regression is proper. Function (6) shows the general model:

$$y = a + b_1 x_1 + b_2 x_2 + b_3 x_3 + ... + b_n x_n \quad (6)$$

where y is the dependent variable, x_1, x_2 … x_n are independent variables, a is a constant, and b1, b2,…bn are the regression coefficients.

3 Drainage Pavement Program

3.1 Introduction to Drainage Pavement

In Taiwan most of the rain falls in the season of summer, which poses a great challenge for asphalt concrete's drainage performance and water resistance. Besides, researches [10] have indicated that since the north-and-south traffic is heavy, to enhance transportability, tire pressure and axle loading often times exceed standard benchmarks so that road surface begins to deteriorate before their expected service life ends.

Porous asphalt concrete is a gap-graded asphalt concrete that contains a high ratio of course materials. In addition to coarse materials, fiber is also a main ingredient of porous asphalt. Fiber can help stabilize the high content of asphalt mastic in the porous asphalt. Mineral fillings, such as lime and fly ash, are used as anti-stripping additives. Porous asphalt concrete also contains same-graded asphalt mastic or modified asphalt [11]. To understand whether porous asphalt concrete can fit the properties of Taiwan's highways, a trial pavement project was made to the 132k+200~132k+400 section on the northbound side of National Highway No. 3 and the performance was observed and recorded before and after the completion of the pavement project.

3.2 Roughness Index

Roughness can serve as an important indicator in determining the road surface's serviceability. The roughness can be measured from three cross-sections: vertical, and horizontal direction. The deformation of the road surface tends to impose an acceleration impact on a vehicle's side and vertical. Vertical acceleration can greatly influence ride-ability and a user's experience in riding on a deformed road surface. Sideward acceleration will cause the vehicle to jolt. The roughness gradually decline with the influence of the repetitive action of car loads, periodical changes in environmental factors, and the increasing age of the servicing road.

When the roughness goes down to a certain point, the road surface will fail to meet the basic needs for readability. At this time, re-pavement or repair should be made to improve the roughness to maintain its serviceability. Different measurement equipment and concepts produce different roughness indices. In recent years, most international highway management agencies have adopted the international roughness index (International Roughness Index, IRI) as the basis for roughness measurement. International roughness index is a standardized roughness indicator and is similar to response-type road roughness measuring systems in that both use the quarter car to

Table 1. Relation between roads, international roughness index, velocity

Roads	IRI (m/km)	Velocity (km/hr)
Runways and highways	0.25~1.75	114~105
New road surface	1.25~3.50	108~98
Old and worn road surface	2.25~5.75	104~86
Regularly maintained unpaved roads	3.25~10.00	98~60
Damaged roads	4.00~11.00	94~57
Uneven unpaved roads	>7.75	<77

move on the road surface cross-section at a specified speed and analyze accumulated vertical displacement of the immediate response by suspension system within the travel distance. The standard car speed is set at 80 km/hr and the results are expressed in Table 1.

4 Analysis and Conclusion

4.1 Data Compilation and Analysis

The interval between the road mark 55.5 km and 59 km on National Highway No. 3 was paved with drainage in October 2004 with eight different designs. In this research, examination was carried out on lane three of the section at an interval of four months to observe the IRI value. Also, how the corresponding designs influenced on the drainage pavement performance was also analyzed as shown in Table 2.

The research uses the drainage pavement on National Highway No. 3 as the specimen for the prediction of roughness. In grey forecast, the four periods of data is input to predict the roughness in March 2006. In genetic programming, likewise, the four

Table 2. Raw data of the drainage pavement roughness(m/km)

	2005.12	2005.4	2005.8	2005.11	2006.3
Cross-section specimen	IRI Average(m/km)				
Factory-mixed 2.5cm	1.25	1.95	2.50	2.33	2.37
Pre-mixed 2.5cm	1.55	2.24	2.27	2.46	2.49
Mineral fiber 3cm	1.44	2.03	2.83	2.43	2.73
Wood fiber 3cm	1.66	2.37	2.90	2.72	2.87
Wood fiber 2.5cm	1.90	2.58	2.24	2.87	3.58
Pre-mixed 4cm	1.75	2.28	2.63	3.74	3.95
Modified type Ⅲ 2.5cm	1.75	1.95	2.50	2.33	2.89
Factory-mixed 4cm	2.18	2.34	2.28	2.63	2.78

periods of data is input, initial cluster is set at 100, the maximum generation is 1000,crossover rate is set at 0.9, and mutation rate is set at 0.01. Discipulus genetic programming software is used for the analysis. And in multiple regressions, SPSS is used to create data. The prediction results are shown in Table 3. It is found that genetic programming can help to build up a roughness prediction mode because genetic programming can change in the parameter combination to deliver most suitable prediction results. The RMSE (Root of Mean Square Error) of different prediction methods is shown in Table 4.

Using genetic programming, we get the formula in below:

$$y = \frac{x_4 + x_1 \times x_4}{(x_2)^4} \tag{7}$$

And using multiple regressions, we get the formula (8):

$$y = -0.2927 + 0.298945 \cdot x_1 + 0.086091 \cdot x_2 + 0.00115 \cdot x_3 + 0.9439 \cdot x_4 \tag{8}$$

Table 3. Different predictions of the drainage pavement roughness (m/km)

Cross-section specimen	Gray forecast	Genetic programming	Multiple regression	Real value
Factory-mixed 2.5cm	2.64	2.53	2.46	2.37
Pre-mixed 2.5cm	2.55	2.61	2.70	2.49
Mineral fiber 3cm	2.88	2.64	2.62	2.73
Wood fiber 3cm	3.02	2.86	2.99	2.87
Wood fiber 2.5cm	2.89	2.99	3.22	3.58
Pre-mixed 4cm	4.72	3.98	3.98	3.95
Modified type Ⅲ 2.5cm	2.64	2.61	2.61	2.89
Factory-mixed 4cm	2.73	2.82	3.06	2.78

Table 4. Different predictions of the drainage pavement roughness (RMSE)

Cross-section specimen	Gray forecast	Genetic programming	Multiple regression
Factory-mixed 2.5cm	0.135	0.08	0.045
Pre-mixed 2.5cm	0.03	0.06	0.105
Mineral fiber 3cm	0.075	0.045	0.055
Wood fiber 3cm	0.075	0.005	0.06
Wood fiber 2.5cm	0.345	0.295	0.18
Pre-mixed 4cm	0.385	0.015	0.015
Modified type Ⅲ 2.5cm	0.125	0.14	0.14
Factory-mixed 4cm	0.025	0.02	0.14

4.2 Discussion

Although genetic programming takes the longest time, genetic programming produces better prediction results than traditional multiple regression and grey forecast do. This is mainly because it takes a large amount of long-term data for traditional prediction methods to carry out analysis. And in this research, there are only eight to ten test time periods and tested cross-section specimens, which does not conform to traditional statistical hypothesis. In addition, the roughness degradation mode is not a linear degradation mode, so it is less suitable to use multiple regression for better predictability. Although grey forecast requires less computation time, its predictability is not as good. And the modeling method can only be used for short-term prediction. Considering the characteristics of different cross-sections, the predictability of grey forecast, in contrast to that of genetic programming, produces higher variability. Moreover, in the modeling process, the average data of each different time spot should be used as the modeling parameters. Hence, a too high variability in the data of each time spot will impact predictability. On the other hand, genetic programming can be used for non-linear prediction and is proved to produce better performance in this research.

5 Conclusion

It is found in this research that genetic programming, as a prediction method, can produce better predictability because it does not set limits to its prediction formula structure. Hence, for a research that does not have a large amount of data, genetic programming can serve as an ideal prediction method. In the future, efforts will be continuously made to collect the IRI value of the section of the road. Also, a degradation mode of the highway drainage pavement will be created, using genetic programming, to predict the actual performance of highway drainage pavement.

References

1. Mrawira, D.M., Wile, C.S.: Roughness Progression Model Based on Historical Data: Case Study from New Brunswick. In: Proceeding of Transportation Research Board, Washington, D.C. (2001)
2. Yu, C.H.: Dynamic Analysis of the Pavement Performance, Graduate Program of the Department of Civil Engineering. National Taiwan University, Taipei (1998)
3. Hung, C.T.: The Study on Establishing the Present Serviceability Index and Predictive Model of Flexible Pavement, Graduate Program of the Department of Civil Engineering. National Central University, Chungli (2000)
4. Shih, K.C., Wu, K.W., Huang, Y.P.: Grey Information Relation Theory. Chuanhua Book, Taipei (1994)
5. Jiang, Z.T.: Design and Application of Fuzzy-Logic-Based Grey Prediction Controller, Graduate Program of the Department of Industrial Technology Education. National Normal University, Taipei (1994)
6. Deng, J.L.: Grey Control System. Huazhong University of Science and Technology Press, Wuhan (1988)

7. Deng, J.L., Kuo, H., Wen, K.L., Chang, T.CH., Chang, W.CH.: Gray Forecast Model and Applications. Gau Lih Book, Taipei (1999)
8. Cramer, N.L.: A Representation for the Adaptive Generation of Simple Sequential Programs. In: John, J. (ed.) Proceedings of an International Conference on Genetic Algorithms and the Applications, Grefenstette, CMU (1985)
9. Koza, J.R.: Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge (1992)
10. Lin, F.Y.: An Analysis on the Deformation Property of Porous Asphalt Concrete, Graduate Program of the Department of Civil Engineering, Tamkang University (2000)
11. Drainage Pavement Technical Guidelines, Japan Road Association, Tokyo, Japan (1996)

New Results on Criteria for Choosing Delay in Strange Attractor Reconstruction

Yang Chen[1] and Kazuyuki Aihara[2,3]

[1] School of Information Science and Engineering, Southeast University, Nanjing 210096, China
cheny@seu.edu.cn
[2] Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-Ku, Tokyo 153-8505, Japan
[3] Aihara Complexity Modelling Project, ERATO, JST, 3-23-5 Uehara, Shibuya-Ku, Tokyo 151-0064, Japan
aihara@sat.t.u-tokyo.ac.jp

Abstract. The first minimum of mutual information (MI) is the most widely used criterion for choosing delay in delay-coordinates reconstruction of chaotic attractors. Recently, we proposed to use the independence measures grid occupancy (GO), quasientropy (QE), and Csiszar's generalized mutual information (GMI) as alternative criteria for choosing delay. In this paper, we study whether these criteria can outperform the MI in certain aspect. We generalize the definition of quality factor (QF), originally defined for QE, to make it also applicable for GO and GMI. We show that GO, QE, and GMI can have better QFs than that of MI, which means that they can have more prominent minima than the MI has. We also verify these results by numerical experiments.

Keywords: Chaos, Delay-coordinate, Mutual information, Quality factor.

1 Introduction

An interesting property of dynamical systems is that their multivariate behavior can be approximately "reconstructed" from observations of just one variable by a method called state space (phase portrait) reconstruction. For chaotic systems, this method is also termed strange attractor reconstruction, which has become a fundamental approach in chaotic time series analysis. Basic reconstruction can be done using delay coordinates [1]. That is, the behavior of the system is reproduced by taking the current observed value $e(t)$ and the observed values after equally spaced time lags to obtain "observations" of multiple variables $\{ r_1(t),\ r_2(t),\ r_3(t),\ ... \} = \{ e(t),\ e(t+\tau),\ e(t+2\tau),\ ... \}$, where τ is the delay to be determined. Fraser and Swinney used the first minimum of mutual information (MI) for choosing delay according to Shaw's suggestion. They proposed a recursive algorithm for computing MI and got good results with it. Their paper [2] has been widely cited till now by researchers in various disciplines since its publication. It seems that using MI to analyze nonlinear time series has become a standard approach. In [3], we proposed to use the independence measures grid occupancy (GO), quasientropy

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 946–953, 2008.
© Springer-Verlag Berlin Heidelberg 2008

(QE), and Csiszar's generalized mutual information (GMI) [4] as alternative criteria for choosing delay. We developed a recursive algorithm for computing GMI, which is a generalization of Fraser and Swinney's algorithm for computing MI.

In this paper, we study whether GO, QE, and GMI can outperform the MI in certain aspect. We generalize the definition of quality factor (QF), originally defined for QE in [5], to make it also applicable for GO and GMI. We show that GO, QE, and GMI can have better QFs than that the MI has, which means that they can have more prominent minima than the MI has. We also verify these results by numerical experiments.

2 Quality Factor (QF) of Quasientropy (QE)

2.1 Quasientropy

Let us consider the problem of how to measure the independence of several variables. Denote by $q_r(\cdot)$ the (cumulative) distribution function of variable r, $q_r(u) = \int_{-\infty}^{u} p_r(v)\,dv$, where $p_r(\cdot)$ is the probability density function of r. Without loss of generality, consider two continuous variables r_1 and r_2. Transform r_1 and r_2 by their respective distribution functions as follows:

$$z_1 \equiv q_{r_1}(r_1), \qquad z_2 \equiv q_{r_2}(r_2). \tag{1}$$

Then, $(z_1, z_2) \in [0,1]\times[0,1]$, and we have [3,5]

Lemma 1: r_1 and r_2 are independent if and only if (z_1, z_2) is uniformly distributed in $[0,1]\times[0,1]$.

Lemma 1 shows that we can measure the independence of r_1 and r_2 by measuring the uniformity of the distribution of (z_1, z_2) in $[0,1]\times[0,1]$. To this end, let us partition the region $[0,1]\times[0,1]$ into an $l\times l$ uniform grid and denote by $\xi(i,j)$ the (i,j)th square in the grid:

$$\xi(i,j) = \left[\frac{i-1}{l}, \frac{i}{l}\right] \times \left[\frac{j-1}{l}, \frac{j}{l}\right], \quad (i,j) \in \{1,\ldots,l\}\times\{1,\ldots,l\}. \tag{2}$$

Denote by $p(i,j)$ the probability that (z_1, z_2) belongs to $\xi(i,j)$, i.e.,

$$p(i,j) = \text{Prob}\{\ (z_1, z_2) \in \xi(i,j)\ \}. \tag{3}$$

Then, the QE for measuring the independence of r_1 and r_2 is defined as

$$\beta(r_1, r_2) = \sum_{i=1}^{l}\sum_{j=1}^{l} f\left(p(i,j)\right), \tag{4}$$

where $f(\cdot)$ is a differentiable strictly convex function on $[0, 1/l]$. Clearly, when $f(u) = u\log u$, QE is nothing but the entropy of *p(i,j)* up to a minus sign.

Due to the Jensen's inequality, we have

$$\beta(r_1, r_2) \geq l^2 f\left(\frac{1}{l^2}\right), \tag{5}$$

with equality if and only if $p(i, j)$ is uniform in $\{1,\ldots,l\}\times\{1,\ldots,l\}$.

If r_1 and r_2 are independent, then (z_1, z_2) is uniform in $[0,1]\times[0,1]$ and, thus, $p(i, j)$ is uniform in $\{1,\ldots,l\}\times\{1,\ldots,l\}$. Then the equality in (5) holds and $\beta(r_1, r_2)$ reaches its minimum $l^2 f\left(1/l^2\right)$. Conversely, if r_1 and r_2 are not independent, then there exists l_0 such that for any $l > l_0$, $\beta(r_1, r_2)$ cannot reach its minimum [5]. Thus, for a large enough l, the minimal $\beta(r_1, r_2)$ implies independent r_1 and r_2.

2.2 Quality Factor (QF) of QE

There are infinitely many strictly convex functions. Let us consider the effect of different convex functions on the performance of QE. In [5], we proved that

$$\beta(r_1, r_2) \leq \left(l^2 - l\right) f(0) + lf\left(\frac{1}{l}\right). \tag{6}$$

In order to better differentiate between independence and "near" independence, it is preferable that QE have high sensitivity around its minimum to the variance of the uniformity of probability density. Thus, we write $p(i, j)$ in the following form

$$p(i, j) = \frac{1}{l^2} + \Delta p(i, j), \quad (i, j) \in \{1,\ldots,l\}\times\{1,\ldots,l\}. \tag{7}$$

Then, the Taylor expansion of (4) gives [5]

$$\beta(r_1, r_2) \approx l^2 f\left(\frac{1}{l^2}\right) + \frac{1}{2} f''\left(\frac{1}{l^2}\right) \sum_{i=1}^{l}\sum_{j=1}^{l}\left[\Delta p(i, j)\right]^2. \tag{8}$$

Based on (5), (6), and (8), we can define the quality factor (QF) of QE as

$$Q_\beta\left(f(u)\right) = \frac{f''\left(\frac{1}{l^2}\right)}{\left(l^2 - l\right) f(0) + lf\left(\frac{1}{l}\right) - l^2 f\left(\frac{1}{l^2}\right)}, \tag{9}$$

where, the numerator is twice the amplification rate of the variance between $p(i, j)$ and the uniform probability distribution in (8), and the denominator is the maximum dynamic range of $\beta(r_1, r_2)$ derived from (5) and (6). The greater the QF, the more sensitive is QE to the change in the uniformity of probability distribution around its minimum and, thus, the sharper is the shape of the minimum of QE.

An illustration of the QF of QE is shown in Fig. 1. Suppose that QE is plotted versus $\sum_{i=1}^{l}\sum_{j=1}^{l}\left[\Delta p(i, j)\right]^2$, and φ is the angle that is tangent to the curve of QE at the minimum of QE, then

$$Q_\beta = \frac{2 \cdot \tan \varphi}{\max \beta - \min \beta}. \tag{10}$$

Clearly, when Q_β is large (small), QE is sharp (blunt) around its minimum.

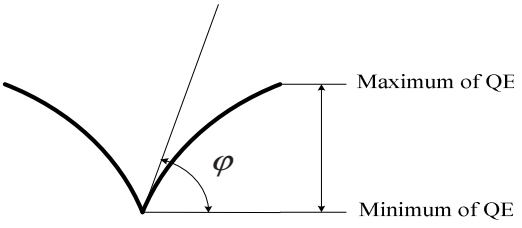


Fig. 1. Illustration of quality factor (QF) of QE

3 QF of Grid Occupancy (GO)

Refs. [3,6] described an interesting phenomenon that independence can be measured by simply counting the number of occupied squares: Given N realizations (observed points) of (r_1, r_2) denoted by $(r_1(t), r_2(t))$, $t = 1, \ldots, N$, where $N > 1$, let

$$z_1(t) = q_{r_1}(r_1(t)), \quad z_2(t) = q_{r_2}(r_2(t)), \quad t = 1, \ldots, N, \tag{11}$$

and let

$$m(i,j) = \begin{cases} 1 & \text{if } \exists t_0,\ 1 \le t_0 \le N \text{ s.t. } (z_1(t_0), z_2(t_0)) \in \xi(i,j), \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

where $(i, j) \in \{1, \ldots, l\} \times \{1, \ldots, l\}$. Clearly, $m(i, j) = 1$ if there is at least one transformed point in square $\xi(i, j)$, or, square $\xi(i, j)$ is occupied. Then, the independence measure GO is defined as minus the ratio of occupied squares, i.e.,

$$\alpha(r_1, r_2) = -\frac{\sum_{i=1}^{l} \sum_{j=1}^{l} m(i, j)}{l^2}. \tag{13}$$

The more the occupied squares, or the less $\alpha(r_1, r_2)$, the more independent are r_1 and r_2 [3,6]. Indeed, we can prove that the expectation of GO is a subclass of QE [7]:

$$E[\alpha(r_1, r_2)] = \beta(r_1, r_2)\Big|_{f(u) = \frac{(1-u)^N - 1}{l^2}}. \tag{14}$$

Therefore, we can define the QF of GO as

$$Q_\alpha = Q_\beta\left(\frac{(1-u)^N - 1}{l^2}\right) = \frac{N(N-1)\left(1-\frac{1}{l^2}\right)^{N-2}}{l\left\{\left(1-\frac{1}{l}\right)^N - 1 + l\left[1-\left(1-\frac{1}{l^2}\right)^N\right]\right\}}. \tag{15}$$

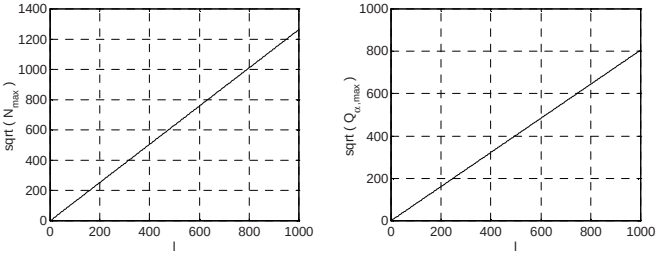


Fig. 2. Numerical study of QF of GO. Left: $\sqrt{N_{\max}}$ versus l. Right: $\sqrt{Q_{\alpha,\max}}$ versus l

Since Q_α is not easy to be analyzed, we study it using numerical methods (see Fig. 2). We vary l from 2 to 1000. For each l, find the N, denoted $N_{\max}$, that maximizes Q_α to $Q_{\alpha,\max}$. In Fig. 2, we can find

$$N_{\max} \doteq \left(\frac{5}{4}l\right)^2, \quad Q_{\alpha,\max} \doteq \left(\frac{4}{5}l\right)^2 = \mathcal{O}\left(l^2\right) \tag{16}$$

4 QF of Generalized Mutual Information (GMI)

Let us now consider the generalized mutual information proposed by Csiszar [4]:

$$\gamma(r_1, r_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{r_1}(u)\, p_{r_2}(v)\, f\left(\frac{p_{r_1 r_2}(u,v)}{p_{r_1}(u)\, p_{r_2}(v)}\right) du dv \tag{17}$$

where $f(\cdot)$ is strictly convex on $[0,\infty)$. Clearly, when $f(u) = u\log u$, GMI reduces to the classical MI

$$I(r_1, r_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{r_1 r_2}(u,v) \log \frac{p_{r_1 r_2}(u,v)}{p_{r_1}(u)\, p_{r_2}(v)}\, du dv, \tag{18}$$

Indeed, we can show that [3,7]

$$\gamma(r_1, r_2) = \int_0^1\int_0^1 f\left(p_{z_1 z_2}(u,v)\right) du dv \tag{19}$$

where z_1 and z_2 are as defined in (1). It is easy to verify that

$$\lim_{l\to\infty}\sum_{i=1}^{l}\sum_{j=1}^{l} f\left(\frac{p(i,j)}{\frac{1}{l}\cdot\frac{1}{l}}\right)\cdot\frac{1}{l}\cdot\frac{1}{l}=\int_0^1\int_0^1 f\left(p_{z_1z_2}(u,v)\right)dudv \tag{20}$$

.

As we are mainly concerned about the situation when l is fairly large, according to (20), we can define the QF of GMI as

$$Q_\gamma\left(f(u)\right)=Q_\beta\left(\frac{f\left(ul^2\right)}{l^2}\right)=\frac{l^3 f''(1)}{(l-1)f(0)+f(l)-lf(1)}. \tag{21}$$

Table 1. QFs of QE and GMI

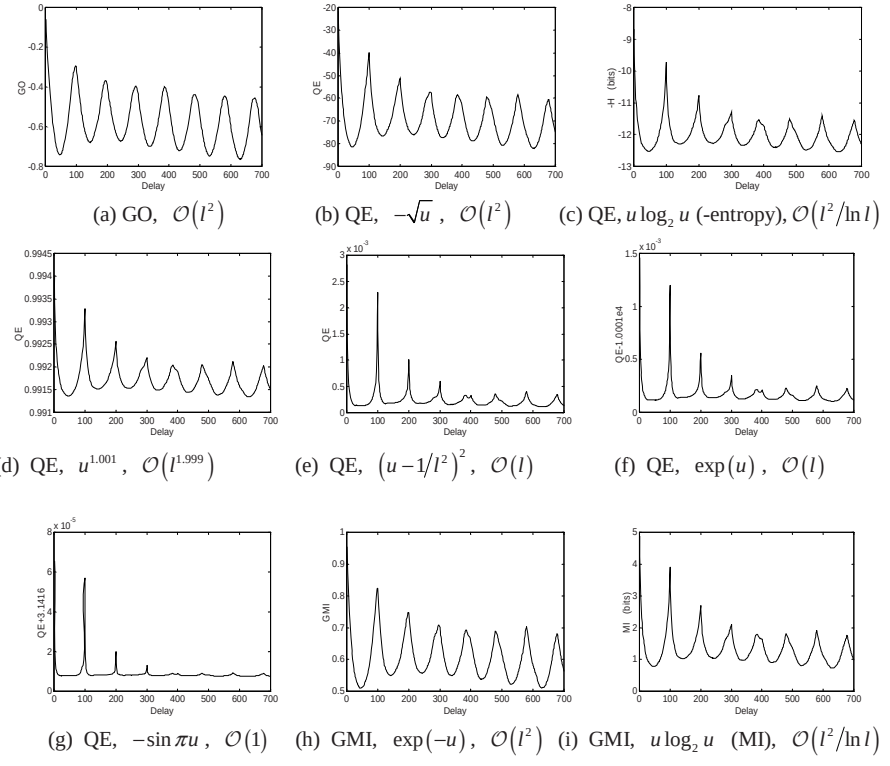
$f(u)$	$Q_\beta\left(f(u)\right)$	$Q_\gamma\left(f(u)\right)$
$-u^a$ $(0<a<1)$	$\frac{a(1-a)l^2}{1-l^{a-1}}=\mathcal{O}\left(l^2\right)$	$\frac{a(1-a)l^2}{1-l^{a-1}}=\mathcal{O}\left(l^2\right)$
$u\log u$	$\frac{l^2}{\ln l}=\mathcal{O}\left(\frac{l^2}{\ln l}\right)$	$\frac{l^2}{\ln l}=\mathcal{O}\left(\frac{l^2}{\ln l}\right)$
u^a $(a>1)$	$\frac{a(a-1)l^2}{l^{a-1}-1}=\mathcal{O}\left(l^{3-a}\right)$	$\frac{a(a-1)l^2}{l^{a-1}-1}=\mathcal{O}\left(l^{3-a}\right)$
a^u $(a>0,\ a\neq 1)$	$\frac{a^{\frac{1}{l^2}}\ln^2 a}{l^2-l+la^{\frac{1}{l}}-l^2a^{\frac{1}{l^2}}}=\mathcal{O}(l)$	$\frac{l^3 a\ln^2 a}{l-1+a^l-al}=\begin{cases}\mathcal{O}\left(l^2\right) & 0<a<1\\ \mathcal{O}\left(\frac{l^3}{a^l}\right) & a>1\end{cases}$

5 Orders of QFs with Respect to *l*

Table 1 shows the Q_β's and Q_γ's of some convex functions and their orders with respect to l as $l\to\infty$. Note the Q_β and Q_γ of $-u^a$ $(0<a<1)$, and the Q_γ of a^u $(0<a<1)$. Their orders are all $\mathcal{O}\left(l^2\right)$, which are higher than $\mathcal{O}\left(l^2/\ln l\right)$, the order of the minus entropy and the MI. For the case of GO, (16) already shows that it can have the QF of $\mathcal{O}\left(l^2\right)$. This means that, when l is large enough, these measures will have sharper minima than the minus entropy and the MI.

6 Numerical Experiments

As in [3], suppose that we have the time evolution $\left(x(t),y(t)\right)$ of Rossler chaotic system. The observed sequence $e(t)$ is obtained by adding noise to $x(t)$. We use $\left(e(t),e(t+\tau)\right)$ to reconstruct the original chaotic attractor, where τ is the delay to be determined. We examine the variation curves of $\alpha\left(e(t),e(t+\tau)\right)$, $\beta\left(e(t),e(t+\tau)\right)$, and $\gamma\left(e(t),e(t+\tau)\right)$ versus τ. Figure 3 shows these curves where some different



(a) GO, $\mathcal{O}(l^2)$ (b) QE, $-\sqrt{u}$, $\mathcal{O}(l^2)$ (c) QE, $u\log_2 u$ (-entropy), $\mathcal{O}(l^2/\ln l)$

(d) QE, $u^{1.001}$, $\mathcal{O}(l^{1.999})$ (e) QE, $(u-1/l^2)^2$, $\mathcal{O}(l)$ (f) QE, $\exp(u)$, $\mathcal{O}(l)$

(g) QE, $-\sin \pi u$, $\mathcal{O}(1)$ (h) GMI, $\exp(-u)$, $\mathcal{O}(l^2)$ (i) GMI, $u\log_2 u$ (MI), $\mathcal{O}(l^2/\ln l)$

Fig. 3. GO, QE, and GMI of delay-coordinate variables of Rossler chaotic system versus delay τ. All plots are calculated using 65536 sample points. $l=100$ is used for GO and QE. The unknown cumulative distribution functions are realized by sorting [5]. GMI is calculated using the recursive algorithm developed in [3]. Delay is measured in sampling number. The convex functions *f*(*u*)'s used in (b)–(i) and the orders of QFs of (a)–(i) are labeled in their captions. The first (leftmost) minima in (a)–(d), (h), and (i) are good choices of the delay for reconstruction.

convex functions are used for QE and GMI. We can see that the order of QF correctly predicts the shape of minima: The higher (lower) the order, the sharper (blunter) the minima. Note that the minima in Figs. 3(a), 3(b), and 3(h) are sharper, and their positions are more definite than those in Figs. 3(c) and 3(i). Namely, these measures outperform the minus entropy and the MI in the prominence of minima.

7 Conclusions

The MI is the most widely used criterion for choosing delay in strange attractor reconstruction. However, there is no evidence to show that it is the optimum criterion. When QE is plotted versus the variance from uniform distribution, the QF is twice the

slope of QE at its minimum over the dynamic range of QE and, therefore, indicates the sharpness of QE around its minimum. The expectation of GO is a subclass of QE. GMI is the limit of a special form of QE as $l \to \infty$. Using these two facts, respectively, we have defined the QFs of GO and GMI. By carefully studying the QFs, we have shown that GO, QE, and GMI can have more prominent minima than the minus entropy and the MI. We have also verified these results by numerical experiments.

Acknowledgments. The authors thank H. Shimokawa for his valuable comments. This work was supported by the National Natural Science Foundation of China (Grant 60202014), Grant-in-Aid for Scientific Research on Priority Areas 17022012 from MEXT of Japan, Industrial Technology Research Grant Program in 2003 from NEDO of Japan, and the Southeast University Excellent Young Teachers Grant.

References

1. Takens, F.: Detecting Strange Attractors in Turbulence. In: Warwick 1980. Lecture Notes in Math., vol. 898, pp. 366–381. Springer, Berlin (1981)
2. Fraser, A.M., Swinney, H.L.: Independent Coordinates for Strange Attractors from Mutual Information. Physical Review A 33(2), 1134–1140 (1986)
3. Chen, Y.: New Methods for Measuring Independence with an Application to Chaotic Systems. In: Proceedings of 2002 International Symposium on Nonlinear Theory and Its Applications (NOLTA 2002), Xi'an, China, October, pp. 523–526 (2002)
4. Csiszar, I.: Information-Type Measures of Difference of Probability Distributions and Indirect Observations. Studia Scientiarum Mathematicarum Hungarica 2, 299–318 (1967)
5. Chen, Y.: Blind Separation Using Convex Functions. IEEE Trans. Signal Processing 53(6), 2027–2035 (2005)
6. Chen, Y., He, Z., Yang, L.: A Novel Grid-Box-Counting Method for Blind Separation of Sources. In: Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, May, pp. 243–246 (2001)
7. Chen, Y.: On Theory and Methods for Blind Information Extraction. Ph.D. dissertation, Southeast University, Nanjing, China (March 2001) (in Chinese)

Application of Data Mining in the Financial Data Forecasting

Jie Wang [1] and Hong Wang[2]

[1] College of Teacher Education, Shanxi Normal University, Linfen, Shanxi 041004, P.R. China
wjlkt@163.com
[2] College of Mathematics and Computer Science,
Shanxi Normal University, Linfen, Shanxi 041004, P.R. China
wangh@sxnu.edu,cn

Abstract. A method of processing financial data based on wavelet transformation is presented. The data of the financial is essentially an unfixed time sequence. Based on the wavelet transform, the series obtained after decomposition contains information. Basically, the wavelet decomposition uses a pair of filters to decompose iteratively the original time series. It results in a hierarchy of new time series that are easier to model and predict. Regarded as a signal, the time sequence is decomposed into different frequency channels (as a filtering step) .These filters must satisfy some constraints such as causality, information lossless, etc. And reconstruction is used to analyze and forecast the time sequence. Examples show that the new method is more effective than the traditional AR model forecast in some aspects.

Keywords: Data mining, Wavelet transform, reconstruction.

1 Introduction

In time series analysis and prediction, there are many methods or technologies which using artificial intelligence, machines study, neural networks, statistical theory, wavelet transform, genetic algorithm, the database technology and the policymaking theory concept, understands and analyzes the complex relations which hidden in the massive data, has the ability of processing different type data. The data mining algorithms have the validity and the probability. The data mining result usually is useful, definite, and can interactively excavate knowledge in the different level from different information source.

The wavelet analysis is a rapid development domain in modern mathematics. At present, it is widely applied in the signal analyzing and imagery processing. But the actually which using the wavelet analysis to analysis and forecast the financial's data are not often. In practical work, forecasting the financial data is extremely important, it plays important role in company's localization and market decision-making. Essentially speaking, the corporate financial data, such as sales volume, profit and disbursement and so on all is one kind of time series. They have the same characteristic as the usual analysis signal, these data can be forecasted through the wavelet analyzing. Moreover, This kind of corporate financial data also has its characteristic, if one kind of time series is steady, then may defer to the traditional method which using

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 954–961, 2008.
© Springer-Verlag Berlin Heidelberg 2008

model such as AR, MA or AR2, MA to forecast usually can achieve the quite good effect. But generally speaking, the corporate financial data stochastic undulation is very big, such as in this article sales volume achievement, the stability is very bad, it belongs to the model a non- steady sequence kind, this time these traditional forecast method effects are not very good Therefore, this article introduces the wavelet analysis method may decompose the signal wavelet to the different frequency channel. Because decomposes after the signal to be more unitary than on the frequency component the primary signal, and the wavelet decomposed to the signal has made smooth processing, so after decomposition, the time series stability is better than primitive time. After carrying wavelet decomposition on the some non- steady time series, it may be treat as the steady time series to process in the approximate significance, like this can use some traditional the forecast method to decompose after the time series to carry on the forecast, Moreover through the example proved this kind of forecast method effect is good.

Financial data analysis data mining system establish in some correlation rationales, like financial knowledge flexible expression, financial information resources appraisal, estimate and optimize information resources choice and so on. This is also the content, which is researched by precisely the modern finance theory. The financial data analysis data mining technology applies to the following domain: (1) Risk appraisal. The insurance is one risk service; the risk appraisal is one important work of insurance company. (2) Financial investment. The typical financial analysis domain has the investment appraisal and the stock transaction market forecasts, which is concerned with a matter development forecast. (3) Cheat recognition. The bank frequently has the cheating behavior, like the malignancy overdraws which brings the heavy loss to the bank.

2 The Wavelet Transform Theory

Usually, the wavelet decomposition and the heavy construction may realize through the Mallat algorithm. Supposes $\{\psi_{j,n,n\in z}\}$ $\{V_j\}_{j\in z}$ is $L^2(R)$ center more than criteria analyzes φ is the criterion function, $\{\psi_{j,n,n\in z}\}$ is the wavelet base, Then influential the solution through the Mallat algorithm:

$$\begin{cases} c_k^{j+1} = \sum_{l\in Z} c_l^j \left\langle \varphi_{j+1,k}(x), \varphi_{j,l}(x) \right\rangle \\ d_k^{j+1} = \sum_{l\in Z} c_l^j \left\langle \psi_{j+1,k}(x), \varphi_{j,l}(x) \right\rangle \end{cases} \tag{1}$$

The heavy construction type is：

$$c_k^j = \sum_{l\in Z} c_l^{j+1} \left\langle \varphi_{j,k}(x), \varphi_{j+1,l}(x) \right\rangle + \sum_{l\in Z} d_l^{j+1} \left\langle \varphi_{j,k}(x), \psi_{j+1,l}(x) \right\rangle \tag{2}$$

Simply written as

$$c_k^{j+1} = \sum_l h_{l-2k} c_l^j, \qquad d_k^{j+1} = \sum_l g_{l-2k} c_l^j$$

$$c_k^{j+1} = \sum_l h_{k-2l}^* c_l^{j+1} + \sum_l g_{k-2l}^* d_l^{j+1}$$

Among them, $\{h_k\}_{k\in z} \in l^2(Z)$ is produced by type $\frac{1}{\sqrt{2}}\varphi\left[\frac{x}{2}\right] = \sum_k h_k^{\varphi}(x-k)$, may regard as the low pass filter coefficient; $\{g_k\}_{k\in z} \in l^2(Z)$ is produce by type $g_k = (-1)^k h_{1-k}^{\varphi}$, May regard as high passes the filter coefficient, c_k^{j+1} is the $\frac{1}{\sqrt{2}}\varphi\left[\frac{x}{2}\right] = \sum_k h_k^{\varphi}(x-k)\, c_l^i$ approximate signal, d_k^{j+1} is the c_l^j detail signal.

Table 1 listed the decision wavelet function quality essential property and some commonly used wavelet function nature.

Table 1. Commonly wavelet function's some nature

Wavelet function	Orthogonal	Tight structure	Structure length	Symmetry
Haar	have	have	1	yes
Daubechies	have	Approximate	2N-1	yes
Symlets	have	Approximate	2N-1	yes
Meyer	have	Limited length	no	yes

3 Time Series Analysis and Prediction

Let $x_1, x_2, ..., x_l$ be a measured time series. The objective is to predict the value of x_{k+p} using all the observations until the instant *k*. For this purpose, a functional relationship that maps the vector [$x_1, x_2, ..., x_l$] and the value x_{k+p} is to be constructed with the principal concern of minimizing the prediction error. The optimal prediction sequence $x'_{p+1}, x'_{p+2}, ...$ minimizes the expectation (or generalization risk)

$$C = \lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T} E\{(x'_{p+1} - x_{p+1})^2 \,|\, x_k, x_{k-1}, ...\} \tag{3}$$

If the time series is random. It is given by

$$x'_{k+p} = E\{x_{k+p} \,|\, x_k, x_{k-1}, ...\} \tag{4}$$

Which cannot be computed since the information about the time series is limited to l measurements. The criterion to be minimized is the following empirical risk:

$$C_{emp} = \frac{1}{l}\sum_{i=1}^{l}(x_i - x_i^{'})^2 \tag{5}$$

All the expectations are omitted if the time series is deterministic. The relationship between $x_{k+p}^{'}$ and the sequence $x_k, x_{k+1}, \ldots$ are supposed to be nonlinear of an unknown nature.

On the other hand, two problems arise. The model's order choice: When the order is small, the training is simple but the information (about the past) is not enough to predict accurately. However, when the order is high, the training is hard and the approximation error decreases slowly. This is the curse of dimensionality.

According to wavelet transform theory, signal $x(n)$'s wavelet series reconstruction can be replaced by the following finite sum.

$$x(n) = \sum_{j=0}^{J-1}\sum_{k\in Z} c_{j,k}\psi_{j,k}(n) \tag{6}$$

Formula (6) means to project the input signal $x(n)$ into corresponding scale's orthogonal subspace so as to recur signals in different distinguishing level.

To make $x_j(n) = \sum_{k\in Z} c_{j,k}\psi_{j,k}(n)$ is $x(n)$'s projective discrete form in wavelet subspace. The purpose of DWTAF is to produce the discrete reconstruction of $x_j(n)$ $(j = 0,1,\cdots J-1)$.

$v_j(n)$ (As is shown in fig.1) is the approximation of projection x_j:

$$v_j(n) = \sum_{k\in Z} \hat{c}_{j,k}\psi_{j,k}(n) \tag{7}$$

Where, $\hat{c}_{j,k}$ is wavelet coefficient $c_{j,k}$'s discrete approximation.

$$\hat{c}_{j,k} = \sum_{l} x(l)\overline{\psi}_{j,k}(l) \tag{8}$$

Formula (8) is put into formula (7), next result will be gained:

$$v_j(n) = \sum_{l} x(l)r_j(l,n) \tag{9}$$

Where $r_j(l,n) = \sum_{k\in Z} \overline{\psi}_{j,k}(l)\psi_{j,k}(n)$.

Formula (9) is $x_j(n)$'s approximation formula. It is the input signal $x(n)$ and filter $r_j(l,n)$'s discrete convolution form. These filters are consisted of wavelet $\psi(t)$'s inflation and convolution after sampling. They are bandpass filters with constant bandwidth/center frequency ratio to realize input signals' multi-distinguishing space reconstruction.

Under the supposition of time-steadiness and orthogonal, Formulas (10) and (11) can be gained:

$$r_j(l,n) = r_j(l-n) \quad , \quad r_j(m) = r_0(2^j m) \tag{10}$$

Then

$$v_j(n) = \sum_l x(l) r_j(l-n) \tag{11}$$

The discussion above is about the signal demonstration pf discrete wavelet transformed. Next the discrete wavelet transforms adaptive LSM algorithm is to be deduced.

If

$$\mathbf{V}(n) = [v_0(n), v_1(n), \cdots, v_{J-1}(n)]^T \tag{12}$$

$$\mathbf{x}(n) = [x(n), x(n-1), x(n-2), \cdots, x(n-N+1)]^T \tag{13}$$

$$[\mathbf{W}]_{jm} = r_j(m), \qquad j = 0,1,\cdots J-1, m = 0,1,\cdots N-1 \tag{14}$$

$$\mathbf{B}(n) = [b_0(n), b_1(n), \cdots, b_{J-1}(n)]^T \tag{15}$$

Then

$$\mathbf{V}(n) = \mathbf{W}\mathbf{x}(n) \tag{16}$$

Where $\mathbf{W}$ is the wavelet transform matrix, its dimension is $J \times N$. $\mathbf{V}(n)$ are the input signals after passing through discrete transform filter.

In the structure of DWTAF, filter $r_j(n)$ $(j = 0,1,\cdots J-1)$ and multi-group delay line coefficient $b_j(n)$ $(j = 0,1,\cdots J-1)$ constitute expected signal $d(n)$'s predictor. Such prediction is based on input $x(n)$'s J continuous linear combination. The adaptive filtering output signals are:

$$y(n) = \mathbf{V}^T(n)\mathbf{B}(n) = \sum_{j=0}^{J-1} v_j(n) b_j(n) = \sum_{j=0}^{J-1}\sum_{i=0}^{N-1} x(n-i) r_j(i) b_j(n) = \sum_{i=0}^{N-1} \alpha_i x(n-i) \tag{17}$$

Where, $\alpha_i = \sum_{j=0}^{J-1} r_j(i) b_j(n)$.

The adaptive filtering error signals are:

$$e(n) = d(n) - y(n) \tag{18}$$

Coefficient update formula is:

$$\mathbf{B}(n+1) = \mathbf{B}(n) + 2\mu e(n)\mathbf{V}(n) \tag{19}$$

The algorithm convergence condition is

$$0 < \mu < \frac{1}{\lambda_{V\max}} \tag{20}$$

To make $\mathrm{R}_{\mathbf{VV}}$ is $\mathbf{V}(n)$'s self- correlation matrix, $\mathbf{R}_{\mathbf{V}d}$ is $\mathbf{V}(n)$ and expected signal $d(n)$'s cross correlation matrix. So $\lambda_{V\max}$in formula (21) is $\mathrm{R}_{\mathbf{VV}}$'s maximum eigen value.

$$\mathbf{B}(n+1) = \mathbf{B}(n) + 2\mu e(n)\mathbf{D}_w^{-1}\mathbf{V}(n) \tag{21}$$

4 Using the Wavelet Analysis to the Financial Data for Forecasting

The main motivation of using the wavelet decomposition is the easy analysis of the obtained series. The corporate financial data can be regards as one kind of time series. It is non-steady time series, then namely $c^0 = [x_1, x_2, ..., x_n]$. Carries on the financial data using the wavelet before the forecast, must carry on wavelet processing first to c^0. The step of the forecasting method of the financial data is explained as following:

1) Selecting suitable wavelet function to carry on the decomposition. In actual forecast process, the data can be dealt with according to the different question choice and different wavelet mother function, at the same time unifies the consideration different wavelet mother function the different characteristic to forecast the value of the influence. Compares various wavelets function processing signal through the analysis, the result and compares with the theory result, determined the wavelet function quality with the error which produces in the processing.

2) Definition the suitable layer of wavelet decomposition. The greatest criterion determination will be advantageous in a group of forecast. The more criterions are big, the more computation workload is also bigger, and the error also can increase. But, the ruler goes past is greatly more advantageous to from the deeper level clear signal trend analysis, it can cause the time series to be steadier in the actual process. When the time series data's quantity which will be forecasted is not very big, the decomposition layer generally is the 3~5 level.

In fact, the trend c_{nk} contains the slowest change. It is practically noise free. The detail series d_{jk} contains the dynamics at a certain intermediate scale. The greater is *j*, the slower are the change. The low detail series may also be corrupted by noise. As a consequence of this property, training an estimator on these time series is simpler than on the original data. However, in low level detail series, the information is totally embedded in noise, one can simply put at zero the corresponding predictions.

3) Decomposing the time series data c^0.

Processing wavelet transform coefficients by formula (6) to (17). Then, the forecasting result is obtained.

5 Application

The series studied here represents the bank savings of a province in 2004, as figure 1. The performance of the forecasting is measured by formula (6) to (17). The wavelet function is 'db4'; the layer of wavelet decomposition is 3. The figure 2 is wavelet decomposition coefficients. The figure3 is result of financial data forecasting.

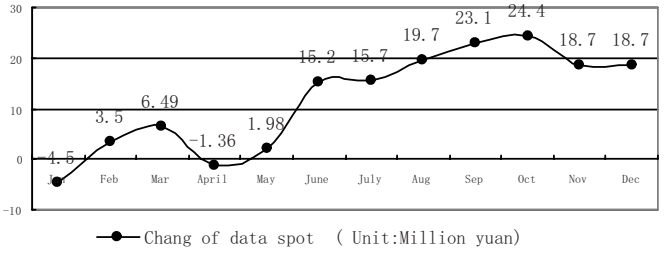


Fig. 1. The bank savings image of a province in 2004

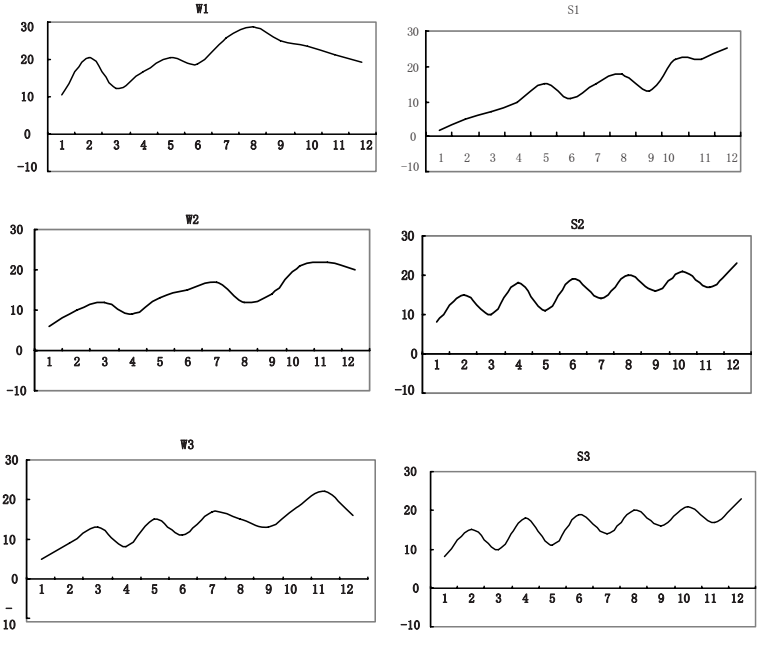


Fig. 2. The high frequency and low frequency of data composition

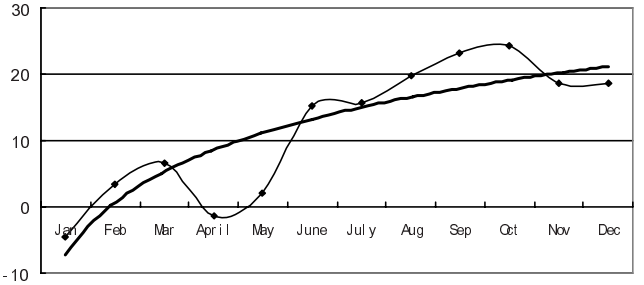


Fig. 3. The result of financial data forecasting (see the thick string)

6 Conclusion

Time series clustering has been shown effective in providing useful information in various domains. There seems to be an increased interest in financial time series clustering as part of the effort in temporal data mining research. It is based on the wavelet transform. The series obtained after decomposition contains information. It may also be used to separate noise from relevant information. Under some conditions, the wavelet decomposition method reduces the empirical risk. The results obtained through the bank saving of a province in 2004 prove the approach effective.

References

1. Casdagli, M.: Nonlinear Prediction of Chaotic Time Series. Physica D 35, 335–356 (1989)
2. Vieu, P.: Order Choice in Nonlinear Autoregressive Models, vol. 26, pp. 304–328 (1995)
3. Li, Z.C., Luo, J.S.: Wavelet Analysis and Its Application. Electronics industry publishing company, Beijing (2003)
4. Zhao, K., Wang, Z.H.: Wavelet Analysis and Its in Analytical Chemistry Application. Geological Press, Beijing (2000)
5. Soltani, S., Boichu, D., Simard, P., Canu, S.: The Long-Term Memory Prediction by Multiscale Decomposition. Signal Processing 80, 2195–2205 (2000)
6. Xu, K., Xu, J.W., Ban, X.J.: Based on Wavelet Certain Non- Steady Time Series Forecast Methods. Electronic journal 29, 566–568 (2001)

Integration of Named Entity Information for Chinese Word Segmentation Based on Maximum Entropy

Ka Seng Leong, Fai Wong, Yiping Li, and Ming Chui Dong

Faculty of Science and Technology of University of Macau, Av. Padre Tomás Pereira, Taipa, Macau, China
{ma56538,derekfw,ypli}@umac.mo

Abstract. Word segmentation is an essential process in Chinese information processing. Although related researches were reported and made progresses, the Unknown Named Entity (UNE) problem in segmentation is not fully solved. This usually degrades the accuracy of segmentation in general. In this paper, a model to identify UNEs for improving the overall performance of the segmentation is presented. In order to capture the NE information, functions of characters or words are defined with tags. In addition, useful surrounding contexts are collected from a corpus and used as features. The model is constructed based on Maximum Entropy to handle the UNE identification as tagging problem. Empirical experiments show that the overall accuracy of the segmentation is improved after integrating the UNE identification module into the word segmenter.

1 Introduction

Unlike English, Portuguese, and etc., the text of Chinese is written in a continuous way without any delimiters to specify the words explicitly. In Chinese information processing, it is well known that the first and essential step is word segmentation. In the literature, although some research work [5], [7] related to word segmentation have been reported and made progresses, the segmentation ambiguities and unknown words problems have not been fully solved. In particular, unknown named entity (UNE) degrades the accuracy of word segmenters a lot.

In our previous research, Maximum Entropy was applied to implement the word segmentation module MESEG (Maximum Entropy Segmentation) [1] because of its flexibility and the feature based model that allows arbitrary features to be integrated into the model. Although MESEG performs much better than our initial segmentation model (CSAT [2]) based on statistics-based approach, it cannot fully resolve the problem caused by unknown words, especially in the case when named entities are frequently occurring in the articles. For the sake of this, we believe that the performance of MESEG can be further improved by integrating the information of NEs.

In this paper, an Unknown Named Entity Identification (UNEI) model based on ME is proposed and integrated into MESEG. The proposed segmentation process is a three-phase analysis, where preliminary segmentation candidates produced by MESEG will be used as the input to the UNEI model. The UNEI model will tag for the segmentation candidates. Then the UNEs will be identified based on the tag information. The tags

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 962–969, 2008.
© Springer-Verlag Berlin Heidelberg 2008

should provide useful internal word characteristics (e.g., person name formation) and the contextual information (e.g., job titles, verbs, etc.) surrounding the potential UNE. With this information, UNEs or nouns can be recalled by re-inspecting the preliminary segmentation candidates to further improve the accuracy of the whole segmentation process.

The paper is organized as follows: the problems in named entity or unknown word identification and the related work will be reviewed and discussed in section 2. Section 3 introduces the general model of ME. Section 4 discusses the modeling of UNEI. Then section 5 gives the features formulation for model. The experiments based on the corpora of the City University of Hong Kong from The Fourth SIGHAN Bakeoff will be discussed in section 6 and followed by a conclusion to end this paper.

2 Named Entity Identification and Related Work

Named Entity (NE) or unknown word identification is one of the sub-problems in Chinese word segmentation. The full task of NE identification consists of labeling the NEs in texts with different types such as person, place or organization names. The difficulties to tackle this problem include: 1) NE identification in Chinese cannot depend on the significant information like capitalization in English. Furthermore, Chinese NE can be composed of any common words, which make it easier to be ambiguous in determining the name. Table 1 gives some ambiguous examples. 2) It is impossible to have a full list of lexicons which can cover all the terms in the real texts or speeches because unknown words keep appearing all the time. In particular, most of the unknown words are named entities and new nouns. 3) Different Chinese NEs have different structures. The structures of abbreviations vary a lot even. Suitable context information and word internal structures are needed to be captured for model. 4) There are few NE-tagged corpora. The algorithm for NE identification should be able to conclude useful information or statistics from the limited resources.

Table 1. Ambiguities related with Chinese NE (ambiguities are in italic)

Intersection with previous word	現任 店長 *爲 何*美珊 (*The shop manager is He Meisan*)
Intersection with succeeding word	李華 *乒* 等 領導 (*Li Huaping etc.*)
Surname with a lexicon word	陳 *光明* (*Chen Guangming*)
Being a lexicon word at the same time	*黎明* (*Li Ming, with another meaning of dawn*)
Ambiguous between person and location name	*林江* 愛 踢 足球 (*Lin Jiang likes playing football*) 我們 來到 *林江* (*We go to Lin Jiang*)
Combining ambiguity	工人 *大勝* 努力 工作 (*Worker Dashen works hard*) 國王隊 *大 勝* (*The Kings win*)

In the literature, there have been a number of methods for unknown word or NE identification. Basically, these can be categorized into the following approaches:

(1). Knowledge-based approaches: They encode the linguistic characteristics of unknown words like NEs into specific rules that will be used as the fundamental knowledge for the identification process. For example, Lv et al.'s work [3] captured the internal structure of unknown words and the context relations to predict the unseen

words. However, the construction of rule set is often complicated and time consuming. It lacks of portability to be applied to different languages. Moreover, it is not easy to encode human knowledge into rules in an effective way.

(2). Corpus-based approaches: They are based on training corpus to capture the required statistic information to estimate the occurrence probabilities of NE. Their independency of language makes them become popular nowadays. The models include class-based language models [4], Hidden Markov Model (HMM) [5], ME [6], etc. For class-based language models, the NEs and contextual information around them are defined with some classes. Then the model is trained based on tagged corpus. After that, it is used to identify NEs in Chinese sentences during the segmentation process. In comparison with knowledge-based approaches, they are more adaptive and robust.

3 Maximum Entropy Model

Among the models in corpus-based approaches, ME is more outstanding. It was first presented by E. T. Jaynes in 1950 and has been applied successfully for different areas in NLP including named entity recognition [6]. ME model is a feature-based, probability model which can include arbitrary number of features (unigram, bigram word features and tag features) into consideration in process that other generative models like N-gram model, HMM cannot. Thus, it was chosen and formulated as the solution to the tagging problem for unknown NE identification as stated. The probability model can be defined over $X \times Y$, where X is the set of possible histories and Y is the set of allowable futures or classes. The conditional probability of the model of a history x and a class y is defined as

$$p_\lambda(y \mid x) = \frac{\prod_i \lambda_i^{f_i(x,y)}}{Z_\lambda(x)} \quad . \tag{1}$$

$$Z_\lambda(x) = \sum_y \prod_i \lambda_i^{f_i(x,y)} \quad . \tag{2}$$

where λ is a parameter which acts as a weight for the feature in the particular history. The equation (1) states that the conditional probability of the class given the history is the product of the weightings of all features which are active under the consideration of (x, y) pair, normalized over the sum of the products of the weightings of all the classes given the history x as the equation (2) above. The normalization constant is determined by requiring that $\Sigma_y p_\lambda(y \mid x) = 1$ for all x.

In ME model, to encode the useful information by helpful feature functions (f) to help predicting the future y by the equation (1) is important. Many problems in NLP can be regarded as linguistic classification problems and most of the ME modelings in NLP are using binary feature functions of the history and future. In this research, we adopt binary feature functions. By the way, another important process in building the model is to find the optimized parameters λ of the conditional probability. Actually, the parameters λ can be chosen to maximize the likelihood of the training data using p:

$$L(p) = \prod_{i=1}^{n} p_\lambda(x_i, y_i) = \prod_{i=1}^{n} \frac{1}{Z_\lambda(x)} \prod_{j=1}^{m} \lambda_j^{f_j(x_i, y_i)} . \quad (3)$$

A number of models can be qualified from (3). But according to the ME principle, the target is to generate a model p with the maximum conditional entropy $H(p)$ below.

$$H(p) = -\sum_{x \in X, y \in Y} p(x, y) \log p(x, y) \text{ where } 0 <= H(p) <= \log |y| . \quad (4)$$

4 UNEI as Tagging Problem

To model the Unknown Named Entity Identification (UNEI) from preliminary word segmentation result as a tagging problem, a manually segmented corpus with NE information is required for the training process of the model. This corpus is first converted into another tagged corpus with some custom defined tags. The custom defined tags form a compact tag set to differentiate word category or character position for used in the tagging process. In this paper, we only focus on discovering the NE that cannot be covered in preliminary word segmentation but their types are not concerned. For experiment purpose, two tag sets are defined as stated in Table 2.

Table 2. Tag sets for UNEI problem: Tag set 1: S-B-M-E scheme, Tag set 2: S-B-I scheme

Word or Character Position or Category	Tag Set 1	Tag Set 2
Non named entity characters or words	OTH	OTH
Single character named entity	S_NE	S_NE
Beginning character of named entity	B_NE	B_NE
Middle character(s) of named entity	M_NE	I_NE
End character of named entity	E_NE	I_NE
The previous character or word of named entity	PRE_NE	PRE_NE
The succeeding character or word of named entity	SUCC_NE	SUCC_NE
The middle character or word of two named entities	MID_NES	MID_NES

After that, context information and features are needed to collect to encode the useful information for the tagging process. With the model trained with suitable context and features, given a preliminary segmented sentence, the model is able to tag the words and the unknown characters with either of these tag sets so that the unknown characters can be recognized as some names or nouns finally.

For an example of UNEI, given "現時 郭 晶 晶 是 中國 的 體壇 明星" (*Guo Jing Jing is the sports star of China now*), assuming "郭晶晶" is the unknown person name, the correct tagged sequence of this sentence according to the tag set 1 in Table 2 will be: 現時/PRE_NE 郭/B_NE 晶/M_NE 晶/E_NE 是/SUCC_NE 中國/OTH 的/OTH 體壇/OTH 明星/OTH. Then it will be converted back to a word sequence according to the tags. The conversion is straightforward. For the sentence tagged with tag set 1 above, when it meets the sequences {B_NE, E_NE} or {B_NE, M_NE (one

or more M_NE), E_NE}, then the corresponding characters of those tags will be considered as NE and grouped together as a word. Similarly, if the sentence above is tagged with tag set 2, when the sequences {B_NE, I_NE (one or more I_NE)} are met, the same action will be done also. Hence the results after conversion of the tagged sentence with tag set 1 or 2 are the same. The final word sequence is "現時 *郭晶晶* 是 中國 的 體壇 明星" where recognized NE is in italic. So now, with the UNEI model added after the rough word segmentation process, the whole word segmentation model framework is as shown in Figure 1.

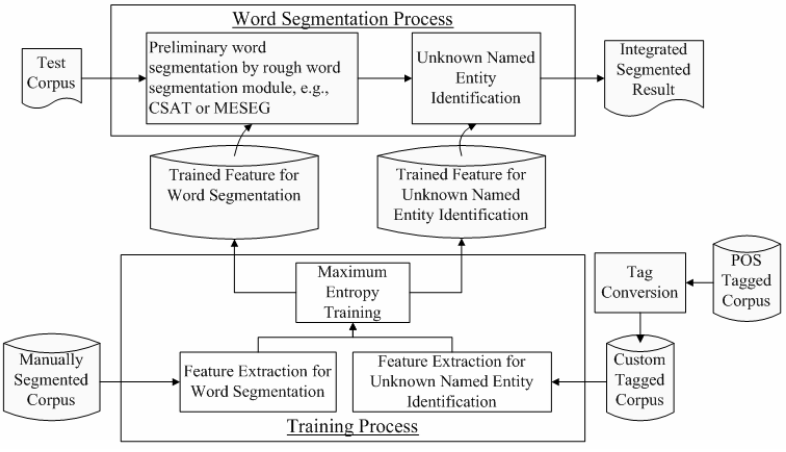


Fig. 1. Framework of ME-based word segmentation model integrated with UNEI process

5 Features Formulation

To construct a model based on ME which is able to perform classification of characters or identification of named entities, features are important. This can be illustrated in this example, "館長 王 景 鴻 編輯 目錄" (*Curator Wong Ching Hung edits the category*). Assume "王景鴻" cannot be recognized after the preliminary word segmentation. But in the sentence, the job title "館長" (*curator*) and the verb "編輯" (*edit*) are surrounding the characters "王 景 鴻". "王", "景", "鴻" are possible component characters of a person name. According to Chinese language, "王", "景" and "鴻" placed in between "館長" and "編輯" can form a person name. All of these context relations and similar information can be collected as some useful features and statistics from a corpus for predicting the unknown named entities. In this section, two feature sets will be described. One set is for MESEG for word segmentation. Another set is for Unknown Named Entity Identification (UNEI) after the preliminary word segmentation.

The following feature templates form the (word segmentation) feature set of MESEG: (WS1): C_n (n = -2 to 2), (WS2): C_nC_{n+1} (n = -2 to 1), (WS3): $C_{-1}C_1$, (WS4): WC_0, (WS5): BT_{-1}, (WS6): $T_{-2}T_{-1}T_0T_1T_2$, (WS7): $Pu(C_0)$. Here C_0 represents the

current character, C_n / C_{-n} represents the character which is at the n^{th} position to the right / left of C_0. Templates WS1~WS3 are character type features. They encode the contexts which are the current character and its surrounding characters. Template WS4 captures the word context in which the current character is a part of that word. Template WS5 is boundary tag feature. It encodes the boundary tag, *BT* (*S*, *B*, *M* or *E* indicate the character position in a word [1]) assigned to the previous character. Template WS6 helps in predicting the pattern of dates and numbers for the word. Four types of label are defined for this feature: "1"- numbers (一, …, 十, 0, …, 9, 零, 百, 千, 萬, 億, 兆), "2"- date characters ("日", "月", "年"), "3"- English characters, "4"- other characters. Finally, template WS7 is to check whether C_0 is a punctuation mark (e.g., ",", "、", "。", etc.). The reason of including this feature is that punctuation marks are good symbols to indicate word boundaries.

The feature templates WS1~WS5 above were used in our previous work [1]. Templates WS6, WS7 are newly introduced in this work. For the example of WS6, consider when a context is "二十年時A", the feature $T_{-2}T_{-1}T_0T_1T_2$ = "11243" is set to 1.

Regarding to the UNEI features, the feature set used should encode the contextual information surrounding the character being concerned. For example, if there is context information that is prefix (e.g., *老*張, *小*李) or postfix (e.g., *張*氏, *李*總) of a name or previous context (e.g., 去, 來到) or succeeding context (e.g., 表示, 報道) of a name, then the character being concerned will be probably a component character of a NE. In the following, the feature set includes: (NE1): W_n (n = -2 to 2), (NE2): W_nW_{n+1} (n = -2 to 1), (NE3): $W_{-1}W_1$, (NE4): T_n (n = -2, -1). Here, W_0 represents the current word, W_n/W_{-n} represents the word which is at the n^{th} position to the right / left of W_0 being considered. *T* represents the custom defined tag in Table 2 assigned to the Chinese word which is at the n^{th} position to the left of W_0.

In applying the feature templates, consider a context with five words assigned with tags of the custom defined tag set 2 in Table 2: "總理/PRE_NE 溫/B_NE 家/I_NE 寶/I_NE 表示/SUCC_NE". Assume "家" is the current character, then e.g., template NE1: W_{-2} = "總理". Template NE2: $W_{-2}W_{-1}$ = "總理溫". Template NE3, $W_{-1}W_1$ = "溫寶". Template NE4, T_{-2} = "PRE_NE", T_{-1} = "B_NE". These features are all set to 1.

6 Experiments and Discussion

To test the performance of the word segmentation with the integration of the UNEI model, experiments were carried out by using the City University of Hong Kong (CITYU) corpora from The Fourth SIGHAN Bakeoff. These include the corpora for word segmentation and POS tagging. The corpus intended for word segmentation are streams of texts where the characters of words are delimited with spaces and marked with boundary tags, while another corpus is labeled with the corresponding POS tags for the words, and is intended for constructing the UNEI model. In preparation for the experiments, the corpora are held out for training and evaluating the models. The evaluation metrics: recall (R), precision (P) and F score (F) in The Fourth SIGHAN Bakeoff were used. The relative data set sizes are presented in Table 3.

Table 3. Data set sizes of the CITYU corpora

Corpus	Training (Words)	Testing (Words)
CITYU1 (Boundary Tags)	1.1M	240K
CITYU2 (POS Tags)	1.1M	180K

The first experiment evaluated the accuracy of the word segmentation module preceding the process of UNEI. In this experiment, CSAT [2] and MESEG [1] were constructed and trained based on the CITYU1 corpus in Table 3. Here, MESEG was trained with two different sets of features in order to determine the best feature set. The first model was constructed with the feature functions: templates WS1~WS5 as stated in section 5. Actually, these are the best features that were concluded in our previous work [1]. For the second set of features, WS6 and WS7 were included to evaluate their effectiveness in the model. Table 4 presents the corresponding scores obtained from the experiments.

Table 4. Scores of CSAT and MESEG on the CITYU corpus of The Fourth SIGHAN Bakeoff

Model	R (%)	P (%)	F (%)
CSAT	91.07	83.01	86.85
MESEG with feature templates WS1~WS5	91.18	88.58	89.86
MESEG with feature templates WS1~WS7	91.61	88.63	90.10

From the empirical results, the F score of CSAT is lower than that of MESEG by a significant amount of 3%. This again illustrates that ME based (MESEG) model is more robust than that of the statistics based (CSAT) segmenter. Moreover, by adding two more features: WS6 and WS7, the F score of MESEG is increased by 0.24%. This shows that the new features WS6 and WS7 do contribute to the segmentation model.

In the second experiment, we tried to evaluate the whole segmentation module by introducing the UNEI model into the process. That is, the best preliminary segmentation candidates produced by MESEG were passed as the input to the UNEI model. Two UNEI models were trained based on the corpora converted from the CITYU2 corpus in Table 3 and feature templates given in section 5. Then two different combinations of model integration were tested, and the effectiveness of the process was evaluated. Table 5 demonstrates evaluations of the different models.

Table 5. Comparison of the segmentation accuracy of MESEG and its integrations with UNEI

Model	R (%)	P (%)	F (%)
MESEG with feature templates WS1~WS7	91.61	88.63	90.10
MESEG with the UNEI model using S-B-M-E scheme tag	90.96	89.38	90.16
MESEG with the UNEI model using S-B-I scheme tag set	91.50	89.57	90.53

From the empirical results, it shows that the overall performance of the segmentation module is improved by integrating the UNEI model into the process. Among the different combinations, the MESEG model integrated with the UNEI model based on the S-B-I scheme outperforms that with the UNEI model based on S-B-M-E scheme.

7 Conclusion

In this paper, the integration of NE information for word segmentation is proposed. The idea was realized by introducing unknown named entity identification (UNEI) model to recognize unknown named entities (NEs) from the preliminary segmentation results. The UNEI model is constructed based on maximum entropy, and the recognition of NEs is treated as tagging problem. Experiments were performed by using the CITYU corpora from The Fourth SIGHAN Bakeoff, and the empirical results show that the proposed model outperforms the previous segmentation modules without UNEI process, with an improvement of 0.43% in its word segmentation precision.

Acknowledgements. The research reported in this paper was partially supported by The Science and Technology Development Fund under grant 041/2005/A and Center of Scientific and Technological Research (University of Macau) under Cativo: 5571.

References

1. Leong, K.S., Wong, F., Li, Y.P., Dong, M.C.: Chinese Word Boundaries Detection Based on Maximum Entropy Model. In: Proceedings of the 11th International Conference on Enhancement and Promotion of Computational Methods in Engineering and Science (EPMESC-XI), Kyoto, Japan (2007)
2. Leong, K.S., Wong, F., Tang, C.W., Dong, M.C.: CSAT: A Chinese Segmentation and Tagging Module Based on the Interpolated Probabilistic Model. In: Proceedings of the 10th International Conference on Enhancement and Promotion of Computational Methods in Engineering and Science (EPMESC-X), Sanya, Hainan, China, pp. 1092–1098 (2006)
3. Lv, Y., Li, T., Yang, M., Yu, H., Li, S.: Leveled Unknown Chinese Words Resolution by Dynamic Programming. Journal of Chinese Information Processing 15(1), 28–33 (2000)
4. Sun, J., Gao, J., Zhang, L., Zhou, M., Huang, C.: Chinese Named Entity Identification Using Class-Based Language Model. In: Proceedings of the 19th International Conference on Computational Linguistics, Taipei, Taiwan, pp. 1–7 (2002)
5. Liu, Q., Zhang, H.P., Yu, H.K., Cheng, X.Q.: Chinese Lexical Analysis Using Cascaded Hidden Markov Model. Journal of Chinese Information Processing 41(8), 1421–1429 (2004)
6. Borthwick, A.: A Maximum Entropy Approach to Named Entity Recognition. PhD thesis, New York University (1999)
7. Xue, N., Shen, L.: Chinese Word Segmentation as LMR Tagging. In: Proceedings of the 2nd SIGHAN Workshop on Chinese Language Processing, Sapporo, Japan, pp. 176–179 (2003)

Evaluating and Comparing Biomedical Term Identification Systems

Yifei Chen, Feng Liu*, and Bernard Manderick

Vrije Universiteit Brussel
Computational Modeling Lab
Pleinlaan 2, 1050 Brussels, Belgium
{yifechen,fengliu,bmanderi}@vub.ac.be
http://como.vub.ac.be/

Abstract. In this paper, we propose a term identification system using conditional random fields (CRFs) on two biomedical datasets. Through employing several sets of experiments, we make a comprehensive investigation for different types of features. The final experimental results reflect that with carefully designed features i.e., adding not only the individual and dynamic features but also the combinational features, our system can identify biomedical terms with fairly high accuracy on both datasets, compared with other top systems already published in the literature.

Keywords: biomedical term identification, named entity recognition, conditional random fields, feature selection.

1 Introduction

Nowadays since the amount of biomedical literature published daily is increasing exponentially, it is not realistic any more to process all the literature manually. Hence literature mining in the biomedical domain emerges more and more important. Biomedical term identification (BTI) is an important and first step. Without finding biomedical terms precisely, it is difficult to associate these terms with corresponding entries in biomedical databases/ontologies, i.e. named entity normalization (NEN), or extract the relations and knowledge from the literature, for example, protein-protein interaction (PPI). However, due to the complex biomedical nomenclature, term identification in biomedicine is often much more difficult than that in newswire domain. For example, biomedical terms can be full names (such as "fatty acid synthase") or abbreviations (such as "FAS"). Even worse, sometimes abbreviations are not acronyms (e.g. a full name "Gamma glutamyl transpeptidase" has an abbreviation "GGTP"). Moreover, a given biomedical term can have several aliases (e.g., a gene for nerve growth factor receptor has at least 10 different names). Furthermore, biomedical terms may have many spelling variations (e.g. "NF-Kappa B" has several variants, "NF Kappa B", "NF KappaB", and "NF-Kappa II").

* Author funded by a doctoral grant of the Vrije Universiteit Brussel.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 970–977, 2008.
© Springer-Verlag Berlin Heidelberg 2008

In order to handle these difficulties, many different approaches have been applied in the past few years, statistical approaches, including Support Vector Machines (SVMs) [1], Conditional Random Fields(CRFs) [2] and so on, dictionary/knowledge-based approaches [3]. Moreover, many systems make use of domain-specific lexicons/ontologies to incorporate domain knowledge to improve their performances further, for example, SWISS-PROT[4], EntrezGene [5] and Hugo Gene Name Consortium (HGNC) [6].

In this paper we will build a BTI system making use of CRFs. Section 2 will describe this system in detail, then Section 3 will apply our system on two biomedical datasets and present the experimental results and finally Section 4 will give a conclusion.

2 Biomedical Term Identification System

2.1 Overview

Figures 1 show our BTI system. During the CRF learning procedure, tokenized and feature-extracted training data is used to build a identification model. And during the CRF classifying procedure, tokenized and feature-extracted test data is put into the trained model, then the outputs are the final identified biomedical terms.

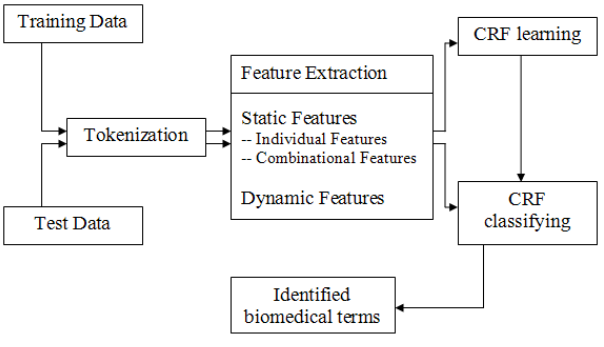


Fig. 1. Biomedical Term Identification System

2.2 Conditional Random Fields

Conditional random fields (CRFs) [8] are a probabilistic framework to segment and label sequence data. The underlying idea is that of defining a conditional probability distribution over label sequences given a particular observation sequence, rather than a joint distribution over both label and observation sequences. The primary advantage of CRFs over hidden Markov models (HMMs) is the ability to relax strong independence assumptions required by HMMs in order to ensure tractable inference. In addition, CRFs avoid the label bias problem, a fundamental limitation exhibited by maximum entropy Markov models

(MEMMs) and other discriminative Markov models based on directed graphical models.

We choose the toolbox Yet Another CRF toolkit (CRF++) [9] to implement our system. CRF++ is a simple, customizable, and open source implementation of CRFs for segmenting/labeling sequential data.

2.3 Data Representation

In this paper we will apply our system on two biomedical datasets. The first dataset *Dataset1* is taken from the data of BioCreAtIvE II challenge [10], Gene Mention Tagging task, which consist of 20000 sentences (15000 sentences as training data and 5000 sentences as test data) derived from MEDLINE abstracts that have been manually annotated for names of genes/proteins.

The second one *Dataset2* is from JNLPBA-2004 Bio-Entity Recognition Task[12]. In this dataset, we not only identify which one is the biomedical term as we do in *Dataset1* but also specify it into one of the following five classes: protein, DNA, RNA, cell line and cell type.

The first pre-processing step is tokenization. Unlike term identification systems in newswire domain, here we can't remove punctuation to make processing straightforward because many biomedical terms contain hyphens, parentheses, brackets and other type of punctuation. Hence we split the sentences based on spaces and punctuation.

Then the second pre-processing step is assigning the class label for each token. We use the traditional BIO representation as follows.

- B: current token is the beginning of a term
- I: current token is inside a term
- O: current token is outside any term

2.4 Feature Extraction

All the features used in our systems for each token are enumerated as follows.

Static Features. Static features are derived from the tokens. Their values are constant as long as the tokens keep unchanged. In total, there are two kinds of static features used in our systems:

- Individual Features: the following individual features are extracted to represent each token.
  - Token: token itself.
  - Orthographic feature: Table 1 shows all the 22 orthographic features used in our system.
  - Prefix and Suffix: bi-, tri- and quad-prefix and suffix of the current token.
  - Lexicon match: Matching biomedical lexicon entries against uni-, bi- and tri-grams (in tokens) starting at the current token. For example, given the sentence "alkaline phosphatase isoenzymes", the uni-gram (the current token) is "alkaline", the bi-gram is "alkaline phosphatase" and the tri-gram is "alkaline phosphatase isoenzymes". There are four features in total, whose values are either 1 if matching or 0 if not matching:

Table 1. Orthographic features

Feature	Example	Feature	Example
DigitNumber	10	LeftParenthesis	(
GreekAlphabet	alpha	RightParenthesis	)
SingleCapital	T	Hyphen	-
InitialCapital	Rho	Backslash	/
LowerMixCapital	GnRH	LeftBracket	[
AlphabetMixDigit	p26	RightBracket	]
AllCapitals	SGPT	Comma	,
AllLowers	stonin	Article	the
Percent	%	Conjunction	and
Period	.	RomanNumber	IV
Colon	:	Others	?,#
SemiColon	;		

  - ∗ Matching gene lexicon entries against uni-gram
  - ∗ Matching protein lexicon entries against uni-gram
  - ∗ Matching protein lexicon entries against bi-gram
  - ∗ Matching protein lexicon entries against tri-gram

  Our biomedical lexicons are refined from SWISS-PROT [4]. After ignoring non-alphabetical letters, digits, roman numbers, Greek letters and some stop words, finally the lexicon contains 68,735 protein names and 31,859 gene names.
- Combinational Features: we also extract the combinational features based on individual features to represent each token.
  - Combination of the current token and its previous token.
  - Combination of the current orthographic feature and its previous orthographic feature.
  - Combination of prefix and suffix features: here we only combine the current tri-prefix and suffix feature and its previous tri-prefix and suffix feature because through our study, it can offer the best result.

Furthermore, by sliding windows, i.e., taking into account a few tokens before and after the current token, we can retrieve the relevant information from the neighboring tokens surrounding the current token.

Dynamic Features. Dynamic features are derived from the predicted class labels. Their values are variable and depend on the running identification system. That is to say, if we use a different system, the predicted class labels maybe are different hence the values of dynamic features are changed. In our systems, we consider the predicted class label of one token preceding the current token as a dynamic feature.

Finally, we will illustrate all the features by Figure 2. The first five columns are individual static features and the following four columns stand for combinational static features mentioned above. The last column "Class label" is the dynamic feature. Suppose we specify the window size as 2 and the current position as 0,

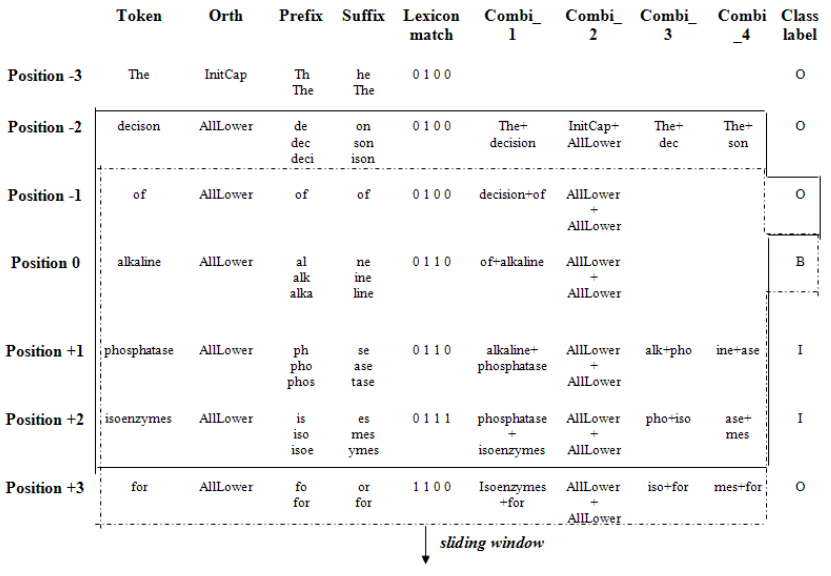


Fig. 2. All the features applied in our systems

hence all the features in Position -2, -1, 0, 1 and 2 are extracted to our systems, which are framed by the solid line in Figure 2. After processing the features in Position 0, we will slide the window to Position 1. Now all the features in Position -1, 0, 1, 2 and 3 are taken out, which are represented by the dashed frame in Figure 2, respectively. Through the sliding window strategy, we can process all the data in our systems.

3 Results and Discussions

3.1 Experimental Purpose

In this paper, our experimental purpose is (1) to find out which types of features can have the positive effect on the performance by setting up several sets of experiments exploring all kinds of features (individual, combinational and dynamic features); (2) to compare the best results achieved with our carefully designed features with other top systems already published in the literature.

3.2 Experimental Preparation

We preform a 5-fold cross validation on training data. For the parameter selection, we try a wide range of cost parameter C from 2^{-10} to 2^{10}. And in order to get the optimal window size in sliding windows, $[-left, right]$ window ranges, $[-3, 3]$, $[-2, 2]$, $[-1, 1]$ and $[0, 0]$ are carried out.

Term identification systems are often evaluated by means of the precision, recall and balanced $F_{\beta=1}$ score, which are defined as follows. "TP", "FP" and "FN" are the numbers of true positives, false positives and false negatives.

$$Precison = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN} \tag{1}$$

$$F_{\beta=1} = \frac{(\beta^2 + 1) \cdot Recall \cdot Precision}{\beta^2 \cdot Recall + Precision} \tag{2}$$

3.3 Experiments with only Individual Features

The effect of individual features in *Dataset1* and *Dataset2* are shown in Table 2. The row of "*Dataset1*" shows the result achieved from *Dataset1*, while the row of "*Dataset2*" shows the result achieved from *Dataset2*. And the result of the row "base system1" is taken from [11], which is in Quartile 1 (TOP 5 of all the systems) in BioCreAtIvE II challenge [10] and makes use of both SVMs and CRFs. Therefore we will use it as a benchmark and compare it with our system. In the same way, the result of the row "base system2" is taken from [12], which is in TOP 3 in JNLPBA workshop and also utilizes CRFs and JNLPBA dataset.

Table 2. Contributions of individual features in *Dataset1* and *Dataset2*

	Precision	Recall	$F_{\beta=1}$
Dataset1	0.7032	0.7588	0.7299
base system1	0.8493	0.8828	0.8657
Dataset2	0.6502	0.6457	0.6480
base system2	0.7030	0.6930	0.6980

From Table 2, it is clear that our system only making use of individual features underperforms the base systems for both datasets. The base system1 can outperform it with $F_{\beta=1}$ score by around 13% while the base system2 can outperform it with $F_{\beta=1}$ score by around 5%. Hence for the next step, we will consider incorporating the dynamic features in our system.

3.4 Experiments by Adding Dynamic Features

We present the results in Table 3, which are obtained by adding dynamic features in our system. Here we only use one dynamic feature, i.e., the predicted class label in Position -1 (assume the current position is 0) illustrated in Figure 2.

Within our expectation, from Table 3, it shows that the dynamic features are critical for our system, which improve $F_{\beta=1}$ score of *Dataset1* by 11.96% and that of *Dataset2* by 5.5%. Now the performance of *Dataset2* has been a little higher than the base system2 while the performance of *Dataset1* is not sufficiently good yet, which is still a little lower than the base system. Therefore, we will continue to add the combinational features into our system.

Table 3. Contributions of dynamic features in *Dataset1* and *Dataset2*

	Precision	Recall	$F_{\beta=1}$
Dataset1	0.8785	0.8271	0.8520
base system1	0.8493	0.8828	0.8657
Dataset2	0.7175	0.6891	0.7030
base system2	0.7030	0.6930	0.6980

3.5 Experiments by Adding Combinational Features

From subsection 2.2, we have known that CRFs can utilize multiple interacting or long-range dependent features. Hence it is worth considering adding combinational features to see if they can have positive impact on the performance. The detailed results are listed in Table 4.

Table 4. Contributions of combinational features in *Dataset1* and *Dataset2*

	Precision	Recall	$F_{\beta=1}$
Dataset1	0.8943	0.8332	0.8628
base system1	0.8493	0.8828	0.8657
Dataset2	0.7249	0.7007	0.7126
base system2	0.7030	0.6930	0.6980

Table 4 shows the detailed results of our system. After adding the combinational features, the performance of *Dataset1* now can be compared to the base system1, since the $F_{\beta=1}$ score, 0.8628, is also in Quartile 1 in [10] and the difference of $F_{\beta=1}$ score between them is not significant at all at 95% confidence interval level. Also we can draw the similar conclusion for *Dataset2*, which increases the $F_{\beta=1}$ score by around 1% and now the difference of $F_{\beta=1}$ scores between them is significant.

4 Conclusion and Future Work

In this paper, we propose a biomedical term identification system using conditional random fields (CRFs) on two biomedical datasets. The final experimental results reflect that with our carefully designed features (individual, dynamic and combinational features), *Dataset1* and *Dataset2* can identify biomedical terms with fairly high accuracy compared with other top systems already published in the literature.

External knowledge resources are a promising point. Some papers [2,11] showed that external lexicons can improve the system greatly. We have used a biomedical lexicon SWISS-PROT [4] in this paper. Next step, we can consider making use of more domain knowledge, e.g., more lexicons (EntrezGene [5], HGNC [6]), full MEDLINE abstracts and web searches.

References

1. Liu, F., Chen, Y., Manderick, B.: Named Entity Recognition in Biomedical Literature using Two-Layer Support Vector Machines. In: The 9th International conference on Enterprise Information Systems (ICEIS 2007), Funchal, Portugal, vol. AIDSS, pp. 39–45 (2007)
2. Klinger, R., Friedrich, C.M., Fluck, J., Hofmann-Apitius, M.: Named Entity Recognition with Combinations of Conditional Random Fields. In: The second BioCreAtIvE Challenge Workshop: Critical Assessment of Information Extractions in Molecular Biology, Madrid, Spain, pp. 89–91 (2007)
3. Humphreys, K., Demetriou, G., Gaizauskas, R.: Two Applications of Information Extraction to Biological Science Journal Articles: Enzyme Interactions and Protein Structures. In: The Pacific Symposium on Biocomputing, vol. 5, pp. 502–513 (2000)
4. Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S., Schneider, M.: The SWISS-PROT Protein Knowledgebases and Its Supplement TeEMBL in 2003. Nucleic Acid Res. 31, D365–D370 (2003)
5. Entrez Gene Database, `http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene`
6. Wain, H.M., et al.: Genew: the Human Gene Nomenclature Database, 2004 updates. Nucleic Acid Res. 32 (Database issue), D255–D257 (2004)
7. Kudo, T., Matsumoto, Y.: Chunking with Support Vector Machines. In: Second Meeting of North American Chapter of the Association for Computational Linguistics (NAACL), pp. 192–199 (2001)
8. Lafferty, J., McCallum, A., Pereira, F.: Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In: The Eighteenth International Conference on Machine Learning (ICML 2001) (2001)
9. CRF++: Yet Another CRF toolkit, `http://crfpp.sourceforge.net/`
10. The second Biocreative - Critical Assessment for Information Extraction in Biology Challenge, `http://biocreative.sourceforge.net`
11. Huang, H., Lin, Y., Lin, K., Kuo, C., Chang, Y., Yang, B., Chung, I., Hsu, C.: High-Recall Gene Mention Recognition by Unification of Multiple Backward Parsing Models. In: The Second BioCreAtIvE Challenge Workshop: Critical Assessment of Information Extraction in Molecular Biology, pp. 109–111 (2007)
12. Collier, N., Kim, J.: Introduction to the Bio-Entity Recognition Task at JNLPBA. In: COLING 2004 International Joint workshop on Natural Language Processing in Biomedicine and its Applications (NLPBA/BioNLP), pp. 73–78 (2004)
13. Chen, Y., Liu, F., Manderick, B.: Improving the Performance of Gene Mention Recognition System using Reformed Lexicon-based Support Vector Machine. In: The 2007 International Conference on Data Mining (DMIN 2007), Las Vegas, Nevada, pp. 228–234 (2007)

Tracking and Visualizing the Changes of Mandarin Emotional Expression

Jun-Heng Yeh, Tsang-Long Pao, Chen-Yu Pai, and Yun-Maw Cheng

Department of Computer Science and Engineering, Tatung University
40 ChungShan North Road, 3rd Section Taipei 104, Taiwan
d9306002@ms.ttu.edu.tw, tlpao@ttu.edu.tw,
g9506017@ms.ttu.edu.tw, kevin@ttu.edu.tw

Abstract. In order to track continuous changes of someone's emotional expressions in Mandarin dialogue, we elaborated an emotion recognition system in continuous Mandarin speech. A new segmentation method was used to estimate the emotional turning points in a dialogue by dividing the utterance into several independent segments, each of which contains a single emotional category. Five basic emotional categories, including anger, happiness, sadness, boredom and neutral, can be recognized in our proposed method. Here, the speech features, MFCC and LPCC, are used in the proposed method. The experimental results show that the new recognition method can provide satisfactory results.

Keywords: emotion recognition, continuous speech, Mandarin.

1 Introduction

Emotion plays a significant role in cognitive psychology, behavioral sciences and human-computer interaction. Humans are especially capable of expressing their feelings by crying, laughing, shouting and subtle characteristics from speech. And classification of emotional states based on the prosody and voice quality requires classifying acoustic features in speech as connected to certain emotional states.

A growing number of research studies in emotion recognition via an isolated short sentence are available to shed some light on the implementations of man-machine interface. However, the short-sentence emotion recognition system may not be able to detect the emotional states correctly because there may have several emotions in a sentence. Even so, human beings speak continuously and people will change emotions when they are triggered by some incidents in the course of speaking. Up to date, few works have focused on automatic emotion tracking of continuous Mandarin speech.

Emotional dimensionality is a simplified description of the basic properties of emotional states. According to the theory developed by Osgood, Suci and Tannenbaum [1] and in subsequent psychological research, the computing of emotions is conceptualized as three major dimensions of connotative meaning: arousal, valence and power. In

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 978–984, 2008.
© Springer-Verlag Berlin Heidelberg 2008

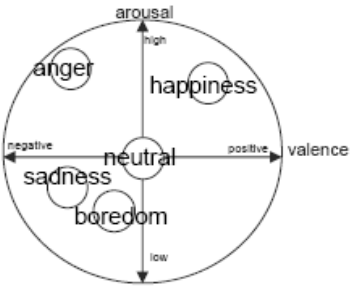


Fig. 1. Graphic representation of the arousal-valence dimension of emotions

general, the arousal and valence dimensions can be used to distinguish most basic emotions. The locations of emotions in the arousal-valence space are shown in Figure 1,which provides a representation that is both simple and capable of conforming to a wide range of emotional applications.

The speech features analysis for emotion recognition [2] has been studied in our previous research to seek for a reliable measurement. These selected speech features are beneficial for us to focus on the emotional segmentation method. Moreover, so far there is relatively little research on the real time emotional speech recognition. Our proposed method tries to implement speech emotion recognition instantly. In our recognition system, we can locate the turning points of emotion changes, and show the emotion categories in a visible interface at the same time.

The rest of the paper is organized as follows: Section 2 describes our features extracted from the speech data and selected from the forward selection method. Section 3 discusses our continuous speech recognition method. Section 4 presents the results of classification experiments. Concluding remarks are given in Section 5.

2 Feature Analysis

In order to find out the optimal combination of all extracted features, we make use of the forward selection to decide the most effective set of features from more than 200 speech features. In the previous research done in our laboratory[3-4], the feature set includes energy, pitch, jitter, shimmer, Formants (F1, F2 and F3), Linear Predictive Coefficients (LPC), Linear Prediction Cepstral Coefficients (LPCC), Mel-Frequency Cepstral Coefficients (MFCC), first derivative of MFCC (dMFCC), second derivative of MFCC (ddMFCC), Log Frequency Power Coefficients (LFPC), Perceptual Linear Prediction (PLP) and RelAtive SpecTrAl PLP (Rasta-PLP). The forward feature selection (FFS) and backward feature selection (BFS) are then used to decrease the dimensions. Finally, MFCC and LPCC are individually obtained from FFS and BFS as the most important features. For speech recognition, MFCC and LPCC are the popular choices as features representing the phonetic content of speech [5]. For each speech frame, 20 MFCCs and 16 LPCCs are used in this system.

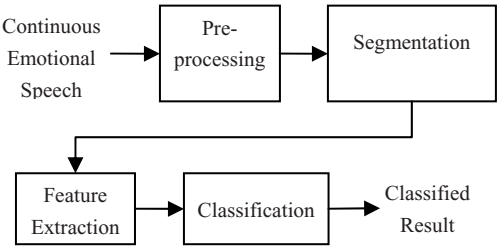


Fig. 2. Block diagram of the emotion recognition system

3 Continuous Emotion Recognition System

Figure 2 shows the block diagram of the emotion recognition system. The analog speech signal is recorded through a microphone and converted to discrete-time signal by sampling and quantization. Framing is used to divide the speech signal into pieces. In this paper, the frame size is set to 256 samples and frame overlapping period is 50% of the frame size. The continuous speech is then divided into suitable length segments. Before further processing, the individual frames are windowed with Hamming window to reduce the discontinuity of both ends of a frame. Finally, these partitions are recognized individually after feature extraction. In previous studies, the methods of continuous segmentation are usually applied in the speech recognition system. To recognize the emotion from continuous emotional speech, the main step is to segment the whole utterance by finding out the turning points of emotional changes. Then, emotion of each segment is recognized in partition basis. The segmentation method used is described as follows. In real-time processing, it is important for the system to be able to detect the endpoints of an utterance so that an assessment can be constructed immediately. There exist some noises in the beginning and the end of the sound. The purpose of endpoint detection is to find the start and the end of meaningful partitions. A simple method to obtain endpoints is to calculate the energy contour and zero-crossing rate contour. The equations for the two energy thresholds and one zero-crossing rate threshold is defined as follows,

$$T_L = \mu_E + \alpha_1 \sigma_E . \tag{1}$$

$$T_U = \mu_E + \alpha_2 \sigma_E , \quad \alpha_1 < \alpha_2 \tag{2}$$

$$T_Z = \mu_Z + \alpha_3 \sigma_Z . \tag{3}$$

where μ_E and σ_E are the corresponding mean and standard deviation of the energy, μ_Z and σ_Z are the corresponding mean and standard deviation of zero crossing rate. The α_1, α_2 and α_3 are parameters, which are obtained by experiments.

Along the sequence of partitions, the first partition with energy greater than T_L is labeled as N_B. If the energies of the next B successive frames are greater than T_L and the

energy of the frame next to B frames is greater than T_L, N_B may be regarded as the beginning of a sound. On the other hand, if the energy of one of the B frames is less than T_L or the energy of the frame next to B frames is less than T_U, it is not the beginning of the sound. In this case N_B will be neglected. After locating the N_B, the next step is to check the zero-crossing rate of all the B frames to see if their zero-crossing rate is greater than T_Z. Now the frame is regarded as the true beginning of the sound, and is labeled as N_S. The frame after N_S with energy greater than T_L means that the sound exists. The first frame after N_S with energy less than T_L is the end of the sound, and is labeled as N_E. As a result, the region of the sound is from N_B to N_E or from N_S to N_E.

After the endpoint detection processing, the number of the segmented partitions may be too large to process. So, it is necessary to reduce the number of partitions. The way is to combine adjacent partitions if they have similar characteristics or too short in length. The mean of the lengths between the beginning and end of the adjacent endpoints are calculated. If the length of the frame is less than the threshold, it will be merged with adjacent frames. The threshold is obtained from experiments. In classification, to let the computer knows what the input signal means, a group of features extracted from frequency components and acoustic models is collected as the basis for the recognition engine. As mentioned in Section 2, the 20 MFCCs and 16 LPCCs features are used to recognize the emotional state of the input utterance by our proposed weighted D-KNN classifier [3].

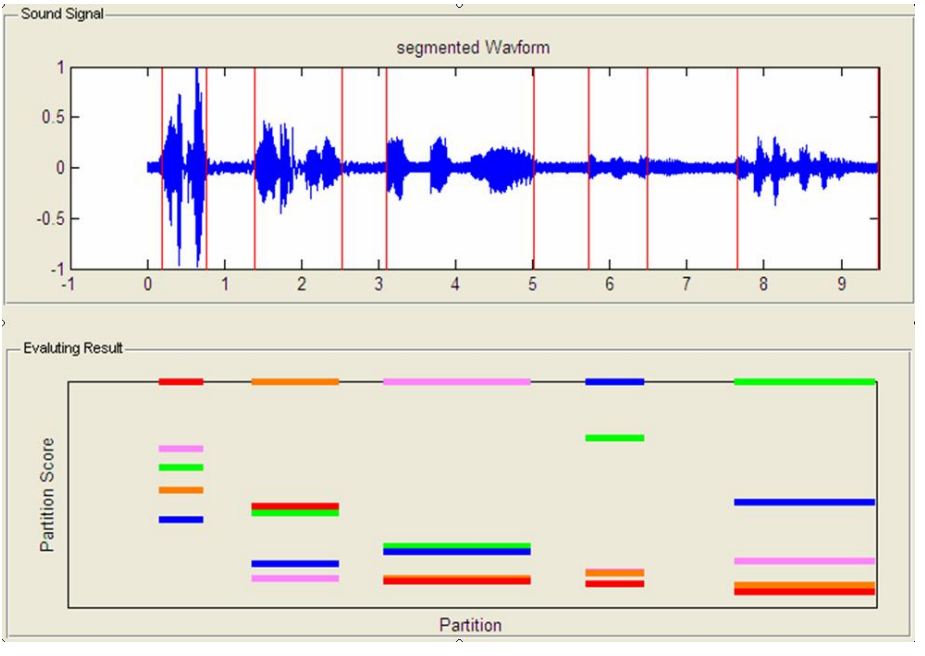


Fig. 3. The interface of the emotion recognition system

From the experimental results done in [6], the Fibonacci weighting function performs best among various weighting schemes. Therefore, we adopted Fibonacci series as the weighting function in this paper.

4 Experimental Results

The objective of our research is to find the turning points of emotional expressions in continuous Mandarin speech. The database in [2] was used. In this paper, utterances in Mandarin were used due to the immediate availability of native speakers of the language. It is easier for speakers to express emotions in their native language than in a foreign language. Twelve native Mandarin language speakers (7 females and 5 males) were asked to generate the emotional utterances. From the corpus, 558 utterances with over 80% human judgment accuracy were selected [2]. The recording format is mono channel PCM with sampling rate of 44.1 kHz and 16-bit resolution. Then we used the features by forward selection mentioned in Section 3 in our proposed recognition method. The testing sentence is a sentence that was combined with emotional short sentences randomly chosen from our database.

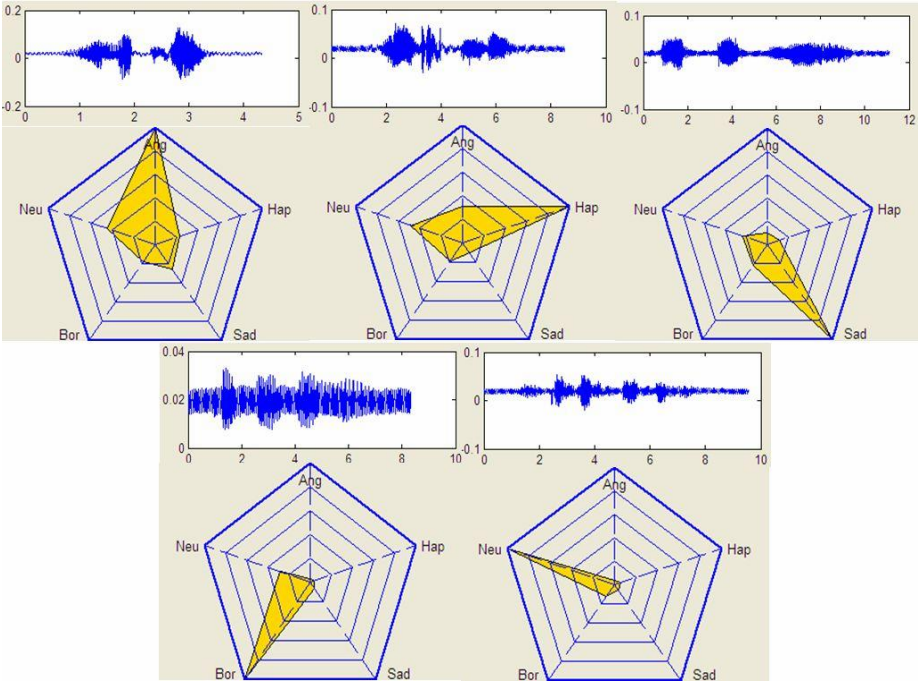


Fig. 4. Radar chart of the emotion recognition system

The interface of our proposed recognition system is shown in Figure 3. The five colors, red, orange, purple, blue and green represent angry, happiness, sadness, boredom and neutral, respectively. In Figure 3, the x-axes of both boxes (Source Signal and

Evaluating Result) are concurrent in time. The Source Signal box shows the results of emotional speech segmentation. In our proposed method, the unit of segmented sentences is an emotional category rather than a word. And the recognition rate achieves 82%. The Evaluating Result shows the relative intensity in each segmented emotional category. Users can easily recognize what is the major or minor emotional expression of a test corpus from the colors.

Each segmented sentences from the concatenated sentences can also be validated individually to estimate the emotional intensity as shown in Figure 4. In this figure, a concatenated sentence with angry, happiness, sadness, boredom and neutral emotions was validated. In our proposed recognition system, each recognized emotional category can be visualized as a "radar chart" to analyze more complicated emotional expressions. It can also present the relative intensity of emotional categories in a segmented sentence as shown in the Evaluating Result box of Figure 3.

In this emotion recognition system, the input speech will be recognized and mapped to the emotional expressions continuously.

5 Conclusion

It is natural for human beings to communicate with others in continuous dialogue. Even though, most proposed emotion recognition methods via voice can only be provided with a fragmented sentence (i.e. a manual and deliberate cutting sentence). To ensure the practicability, the purpose of this paper attempts to address these areas concerned on processing speech signals rather than interpreting the lexicons of speech. Moreover, the benefit from the processing can also tack the change of emotional expression in a signal sentence. The experimental results show that the proposed recognition method can be practically implemented and provide satisfactory results. In our recognition system, we can find out the turning points of emotional changes, and show the emotional categories in a visible interface instantly.

In the growing range of interactive interfaces, the research of emotional voice is still at an early stage, not to mention a paucity of literature on real applications. The crucial difficulty of this subject is how to blend the knowledge of interdisciplinary, especially in speech processing, applied psychology and human-computer interface. There are lots of applications that can benefit from the emotion speech recognition. First, in the hearing-impaired learning system, we can teach the hearing-impaired people to speak more naturally. Second, in the call-center application, it is possible to pick up the angry customer immediately. Third, in the slogan validator, the system may provide a convenient testing tool for marketers and researchers to evaluate a brand slogan that gives consumers the intended impression of the product.

References

1. Osgood, C.E., Suci, J.G., Tannenbaum, P.H.: The Measurement of Meaning. The University of Illinois Press, Urbana (1957)
2. Pao, T.L., Chen, Y.T., Yeh, J.H.: Emotion Recognition from Mandarin Speech Signals. Chinese Spoken Language Processing. In: International Symposium on Digital Object Identifier, pp. 301–304 (2004)

3. Vidrascu, L., Devillers, L.: Annotation and Detection of Blended Emotions in Real Human-Human Dialogs Recorded in a Call Center. In: IEEE International Conference on Multimedia and Expo., pp. 944–947 (2005)
4. Cheng, Y.M., Kuo, Y.S., Yeh, J.H., Chen, Y.T., Pao, T.L., Chien, S.C.: Using Recognition of Emotions in Speech to Better Understand Brand Slogan. In: Proceedings of the International Workshop on Multimedia Signal Processing (2006)
5. Kondoz, A.M.: Digital Speech: Coding for Low Bit Rate Communication Systems. John Wiley & Sons, Chichester (1994)
6. Chang, Y.H.: Emotion Recognition and Evaluation of Mandarin Speech Using Weighted D-KNN Classification. Master Thesis, Tatung University, Taipei (2005)

Exploiting Attribute-Wise Distribution of Keywords and Category Dependent Attributes for E-Catalog Classification*

Young-gon Kim[1], Taehee Lee[2], Sang-goo Lee[3], and Jong-Heung Park[1]

[1] Postal Technology Research Center
Electronics and Telecommunications Research Institute,
Daejeon 305-700, Republic of Korea
{ygkim83,jpark}@etri.re.kr
[2] Electrical and Computer Engineering Department
Carnegie Mellon University, Pittsburgh, PA
thlee94@cmu.edu
[3] School of Computer Science and Engineering / Center for E-Business Research
Seoul National University,
Seoul 151-742, Republic of Korea
sglee@europa.snu.ac.kr

Abstract. E-catalogs are semi-structured documents that consist of multiple attributes and values. Although the conventional text classification techniques are applicable to the e-catalog classification as well, they cannot use the attribute information effectively to improve the classification accuracy. In this paper, we propose an e-catalog classification algorithm by extending Naïve Bayesian Classifier to use the attribute information. Specifically, we focus on exploiting two e-catalog specific characteristics: the attribute-wise keyword distribution and the category dependent attributes. Experiments on real data validate the proposed method.

Keywords: E-catalogs, Attribute information, Classification.

1 Introduction

E-catalogs are the electronic documents that describe the products and services available in e-business environment. All e-business service providers and online shopping malls maintain e-catalogs of available products and services for customers to browse and find what they want. For efficient maintenance, e-catalogs are classified into a specific classification scheme. Most commonly used classification schemes are UNSPSC, eCl@ss and eOTD, which define a hierarchical structure of thousands of

* This work was supported by the Postal Technology R&D program of MKE/IITA. [2006-X-001-02, Development of Real-time Postal Logistics System].

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 985 – 992, 2008.
© Springer-Verlag Berlin Heidelberg 2008

categories. Every new product is registered as an e-catalog after it is classified to an appropriate leaf category.

Although there is no single standard for modeling e-catalogs yet, it is widely accepted that e-catalogs consist of a set of attribute-value pairs. Fig. 1 shows an example of e-catalog that belongs to the category 43171801, 'Laptop Computers'. It has attributes such as 'Product name', 'Company Name' and 'Model Number', and associated values for each attribute. In this paper, we propose an e-catalog classification algorithm that exploits the attribute information. Conventional text classification algorithms can not use the attribute information properly. Specifically, our classification algorithm extends the Naïve Bayesian Classifier to use two e-catalog specific properties:

First, we use the attribute-wise keyword distribution. In [2], which is our previous work, we suggested a modified Naïve Bayes Classifier which copes with the attribute-wise distribution of keywords. Based on this, this paper suggests a new way of assigning weights to attributes so that more discriminative attributes have greater influence on the classification.

Product Name	:	Xnote LM50-DMP2
Company Name	:	LG IBM
Product Description	:	Compact size and slim-design
Model Number	:	LM50-DMP2
Class Code	:	*43171801*
Weight	:	2.7Kg
Memory	:	1GB
...	:	...
Battery Capacity	:	57.72Wh

Fig. 1. An example e-catalog

Second, we use the category dependent attributes. There are two kinds of attributes in e-catalogs, common attributes and category dependent attributes.[3] (also referred to as class specific properties and generic properties)[4] In Fig. 1, Product Name, Company Name, Product Description, Model Number, Class Code are common attributes and Weight, Memory, Battery Capacity are category dependent attributes. While common attributes are used in all e-catalogs regardless of the category they belong to, category dependent attributes are used only in the appropriate categories. For example, Battery Capacity may not be used to describe a product that belongs to 'LCD Monitors' or 'Desktop Computers'. Therefore, the fact that any category dependent attributes exist in an input catalog can be effectively used to the classification. We suggest a way of utilizing this information, by creating another common attribute which contains a list of category dependent attribute names.

The remainder of this paper is organized as follows. In section 2, we investigate the related work. In section 3, we present how we make a structure aware classifier based on the Naïve Bayes Classifier**.** In section 4, we propose a way of utilizing category

dependent attributes of e-catalogs. In section 5, we present experimental results and in section 6, we conclude this paper.

2 Related Work

In [1], three text classification techniques such as Vector Space Model, K-nearest neighbor and Naïve Bayes Classifier were applied to classify e-catalogs and the Naïve Bayes Classifier showed the best accuracy in their experimental result. Since they applied text-classification techniques to e-catalog classification, the attribute information was ignored.

Each of [5],[6], and [7] presents its own way of implementing a structure-aware classifier for semi-structured documents such as HTML, XML. They show that applying attribute-wise distribution of keywords can improve the classification accuracy. [7], the most recent work, takes a similar approach to ours in the sense that they try to manipulate the effect of each attribute by using meta-classifier. However, all of them consider the input documents, those in both training set and test set, have the same set of attributes.

3 Exploiting Attribute-Wise Distribution of Keywords

In this section, we present a way of making a structure-aware classifier, which deals with attribute-wise distribution of keywords. Our scope is limited to common attributes within this section.

3.1 Modified Naïve Bayes Classifier for E-Catalog Classification

In this subsection, we briefly introduce our previous work. The specifics including implementation issues can be found in [2]. Our work starts from the Naïve Bayes Classifier, which uses the formula below to classify an instance that consists of a set of attribute-value pairs.

$$c_{NB} = \arg\max_{c_j \in C} P(c_j) \prod_i P(a_i = v_i \mid c_j) \tag{1}$$

In (1), c_j denotes a category and a_i denotes an attribute that an input data has. The Naïve Bayes classifier picks a category C_{NB} that maximizes the probability of being classified to the category given the value v_i for each attribute a_i. When Naïve Bayes Classifier is applied to the text classification, each word position of an input text constitutes an attribute of the Naïve Bayes Classifier and the probability of having v_i in the text is calculated by counting the occurrences of v_i in the entire texts in the category.[8, 9] In contrast, our model starts from the original Naïve Bayes Classifier rather then the text classifier. That is, we use attributes of the e-catalog such as Product Name and Company Name directly as attributes of the Naïve Bayes Classifier. In Fig. 2, (a) shows the Naïve Bayes Classifier built up based on the structure of the e-catalog in Fig.1.

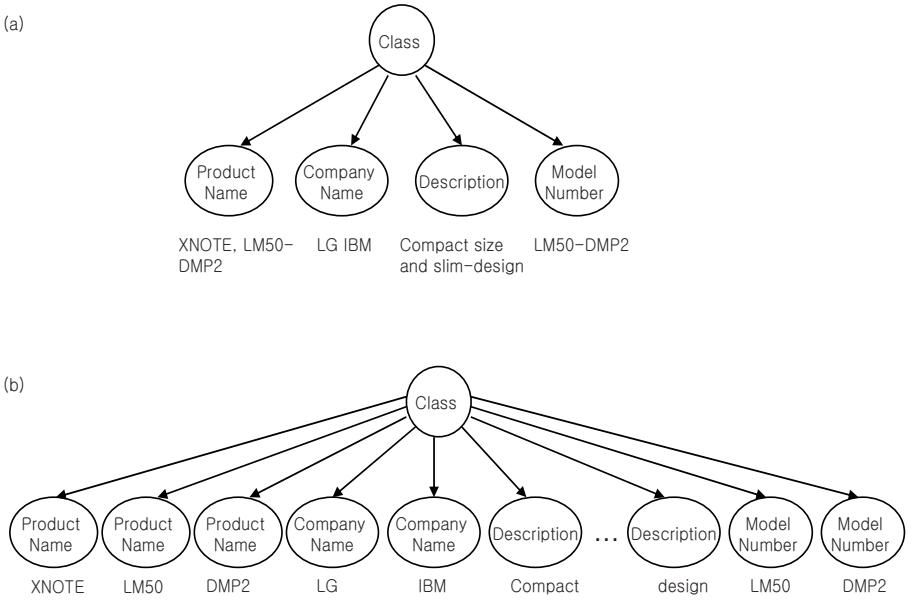


Fig. 2. The structures of Naïve Bayes Classifiers. Texts below the nodes represent the values of an input catalog. (a) is the structure before extension and (b) is the structure after extension.

We can further extend this model to consider the partial matches between attribute values. It is reasonable that the category with more catalogs whose product names have 'XNOTE' is preferred than those with less catalogs even if the product names were not exactly 'XNOTE-LM50'. To make it possible, we parse the value of an attribute into broken terms, and extend the attributes to have multiple values as depicted in the (b) of Fig. 2.

$$v_i = \{t_{i1}, t_{i2}, \ldots, t_{im}\} \tag{2}$$

$$c_{NB} = \arg\max_{c_j \in C} P(c_j) \prod_i \left(\prod_k P(t_{ik} \ appears \ \ in \ a_i \mid c_j) \right) \tag{3}$$

Applying Eq. (2) to Eq. (1) makes Eq. (3). To compute Eq. (3), the frequencies of all terms in each (category, attribute) pair are counted and stored in the training phase. The problem of using Eq. (3) is that longer attributes dominate the classification. The attribute which generates the most terms from the input catalog will have the most influence on the classification, because it will generate the most number of $P(t_{ik} \ appears \ in \ a_i \mid c_j)$ in Eq. (3). However, we want to give higher priority to attributes that are more discriminative in classifying catalogs. We achieve this with the following formula.

$$c_{NB} = \arg\max_{c_j \in C} P(c_j) \prod_i \left(\prod_k P(t_{ik} \ appears \ \ in \ a_i \mid c_j) \right)^{\frac{w_i}{|v_i|}} \tag{4}$$

In this formula, $|v_i|$ is the number of terms parsed out of each attribute of the input catalog, so $1/|v_i|$ removes (normalizes) the effect of the attribute length. w_i is the weight to an attribute that we can assign according to our weighting policy.

3.2 Weighting Policy

In [2], we had to repeat a number of test classifications on the test set to make a weighting policy that gives the best performance. The result of this time-consuming procedure worked well with our data but it may not be possible in other sets of data. It would be more difficult to go through the same weight-search procedure if the new catalogs have higher number of attributes. To solve this problem, we suggest a way of deriving weights from the training data in a single step. Our idea is based on the assumption that the discriminative power of each attribute will follow the discriminative power of the terms contained in the attribute.

Firstly we design a measure of discriminative power of terms. We assume that the fewer categories a certain term occurs in, the more information the term has. For example, the term 'XNOTE' which occurs in only one category is more important than the term 'compact' which occurs in 20 categories. Similar assumption is made for the IDF (Inverse Document Frequency) of TF-[IDF], which is one of the basic concepts in the field of information retrieval.[10] Category Frequency is defined as follows.

Definition 1. Category Frequency (CF)
CF(t) = the number of categories that contain the term t

The terms with lower CF values are more discriminative. Consequently, the attributes whose terms have lower CF values should get higher weights. Our weighting policy is determined as follows.

$$w_i = \frac{k}{average \quad CF \quad of \quad all \quad terms \quad in \quad attribute \quad a_i} \tag{5}$$

Where, k is a constant. Assigning 10 to k, we obtain 5.1 for Product Name, 2.5 for Company Name, 1.1 for Product Description and 5.4 for Model Number, which are similar to the weights gained in [2]. Other functions can be used instead of the one in (5), if they give smaller weights to the attributes with high CF terms.

4 Exploiting Category Dependent Attributes

The category dependent attributes have n-to-m category-attribute relationships to the categories. That is, each category dependent attribute can be used by multiple categories, and a category has multiple category dependent attributes. For example, in G2B classification scheme[11] which is the UNSPSC based classification scheme that our test data follows, each category uses 8.85 category dependent attributes and each category dependent attribute is used in 4.24 categories on average.

The classifier described in the previous section cannot be applied to category dependent attributes since it uses the values from pre-determined set of common attributes. For category dependent attributes, regardless of the values, the existence

itself can be useful information. For example, the information that the input catalog contains Battery Capacity and Main Memory as its category dependent attributes makes it more likely that it belongs to Laptop Computers.

Fig. 2 shows how we use the existence information of category dependent attributes. All of the category dependent attributes are replaced by a new common attribute, 'Attribute Names'. The names of category dependent attributes are concatenated to make a value of the new attribute. Using the category-attribute relationship defined in the classification scheme, the catalogs in the training data are also transformed in the same way before the training phase. Then all the steps described in the previous section can be applied to the new attribute.

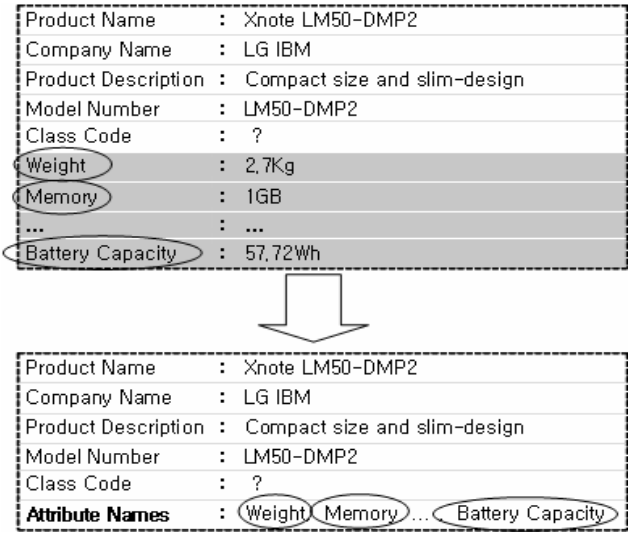


Fig. 3. Making a new common attribute from the names of category dependent attributes

In fact, the exact names of category dependent attributes cannot be known in advance because the category dependent attributes are defined according to the classification scheme. But even the partial match of input attributes with the exact attribute names can be helpful in the classification. The effect of knowing the names of a certain number of category dependent attributes before classification is analyzed through an experiment in the next section.

5 Experimental Results

The e-catalogs used for our experiment are from a product database of Public Procurement Services (PPS) of Korea, which is responsible for procurement for government and public agencies.[11] Our data consists of 495,653 catalogs that are

registered in PPS database until 2006 and the catalogs are classified into 7960 categories according to the G2B classification scheme. The G2B classification scheme is designed for classifying products and services dealt with PPS, based on UNSPSC.

To examine the accuracy of our classifier, we have generated 2 datasets, Data 1 and Data 2. In Data 1, we divided the catalogs into a training set and a test set which are 70% and 30% of the whole respectively. In Data 2, the whole product data is used as a training set and 10,000 randomly sampled product data from the user request log were used as a test set. The user request log contains raw product descriptions that customers submitted for cataloging or purchasing. Through the experiments on Data 2, the performance of our classifier in a real environment can be estimated. The result of the accuracy test on two datasets is shown in Fig. 4. The accuracy is measured in 3 different settings.

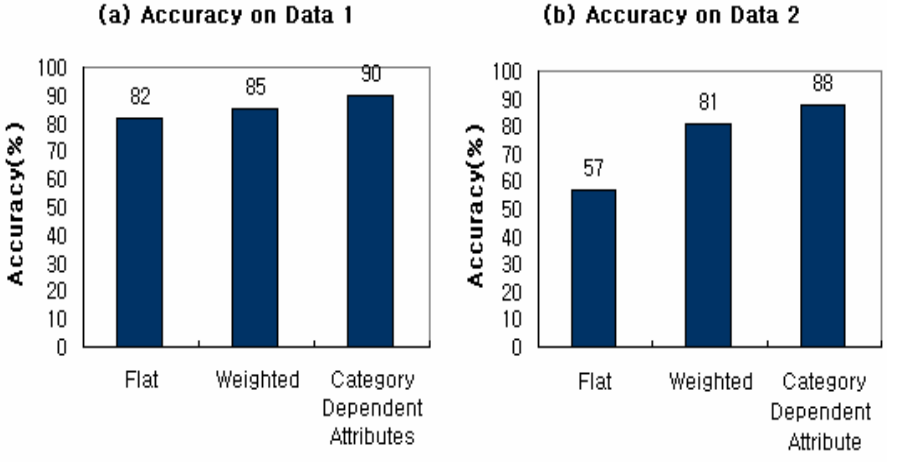


Fig. 4. The result of the accuracy test in two datasets

- **Flat:** Naive Bayes Classifier for flat text classification is directly used
- **Structured-Weighted:** Attribute-wise distribution is used and weights are given according to the discriminative power (Product Name : 5.1, Company Name : 2.5, Product Description : 1.1., Model Number : 5.4)
- **Category Dependent Attributes:** 3 proper category dependent attributes are randomly selected and provided to the input catalog. The weight given to the new attribute is 1.6.

We consider the accuracy of only the first ranked category though our classifier gives top-k categories. The overall accuracy is higher in Data 1 since the input catalogs from data 2 are rather noisy and new to our classifier. However, when all the techniques are used, the accuracy of data 2 is 88% which is close to the 90% accuracy of data 1. This shows that our techniques suggested in this paper have the effect of reducing the noise and can be applied to the real environment.

6 Conclusion

In this paper, we describe two e-catalog specific characteristics and propose ways of utilizing them to improve the classification accuracy. First, considering an e-catalog as a structured document that consists of a set of attribute-value pairs, we suggest a modified Naïve Bayes Classifier that exploits the attribute-wise distribution of keywords. Specifically, we suggest a weighting policy so that more discriminative attributes have larger effects on the classification. As a second characteristic of e-catalogs, we explain the category dependent attributes and suggest a method for utilizing the existence information of category dependent attributes to improve the classification accuracy. Experiments showed that both of the techniques were effective in improving the classification accuracy.

References

1. Ding, Y., Korotkiy, M., Omelayenko, B., Kartseva, V., Zykov, V., Klein, M., Schulten, E., Fensel, D.: GoldenBullet: Automated Classification of Product Data in E-commerce. Business Information System (2002)
2. Kim, Y., Lee, T., Chun, J., Lee, S.: Modified Naïve Bayes Classifier for E-Catalog Classification. In: Lee, J., Shim, J., Lee, S.-g., Bussler, C.J., Shim, S. (eds.) DEECS 2006. LNCS, vol. 4055, pp. 246–257. Springer, Heidelberg (2006)
3. Kim, D., Lee, S.: Catalog Management in e-Commerce Systems. In: Computer Science and Technology (CST 2003) (2003)
4. Hepp, M., Leukel, J., Schmitz, V.: A Quantitative Analysis of Product Categorization Standards: eCl@ss, UNSPSC, eOTD, and RNTD. Knowledge and Information Systems (KAIS) 13(1), 77–114 (2007)
5. Yi, J., Sundaresan, N.: A Classifier for Semi-Structured Documents. In: 6th ACM SIGKDD, pp. 340–344 (2000)
6. Denoyer, L., Gallinari, P.: Bayesian Network Model for Semi-structured Document Classification. Information Processing & Management 40(5), 807–827 (2004)
7. Bratko, A., Filipic, B.: Exploiting Structural Information for Semi-structured Document Categorization. Information Processing & Management 42(3), 679–694 (2006)
8. Mitchell, T.: Machine Learning. McGraw-Hill, New York (1997)
9. Sebastiani, F.: Machine Learning in Automated Text Categorization. ACM Computing Surveys 34(1), 1–47 (2002)
10. Singhal, A.: Modern Information Retrieval: A Brief Overview. Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 24(4), 35–43 (2001)
11. KOCIS (Korea Ontology-based e-Catalog Information System) (Accessed on, April 2008), `http://www.g2b.go.kr:8100/index.jsp`

Learning MultiLinguistic Knowledge for Opinion Analysis

Ruifeng Xu[1,2], Kam-Fai Wong[1], Qin Lu[2], and Yunqing Xia[3]

[1] Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong
[2] Department of Computing, The Hong Kong Polytechnic University, Hong Kong
[3] Institute of Computing Science and Engineering, Tsinghua Universtiy, China
{rfxu,kfwong}@se.cuhk.edu.hk, csluqin@comp.polyu.edu.hk, yqxia@tsinghua.edu.cn

Abstract. Most existing opinion analysis techniques used word-level sentiment knowledge but lack the learning capacity on the behaviors of context-dependent opinion words. Meanwhile, the use of collocation-level sentiment knowledge is not well studied. This paper presents an opinion analysis system, namely *OA*, which incorporates the word-level and collocation-level sentiment knowledge. Based on the observation on the NTCIR-6 opinion training corpus, some word-level and collocation-level linguistic clues for opinion analysis are discovered. Learning techniques are developed to learn the features corresponding to these discovered clues. These features are in turn incorporated into a classifier based on support vector machine to identify opinionated sentences and determine their polarities from running text. Evaluations on NTCIR-6 opinion testing dataset show that *OA* achieved promising overall performance.

Keywords: Opinion Analysis, Linguistic Knowledge Learning, Collocation.

1 Introduction

Aiming at identifying and analyzing the opinions in text, opinion analysis becomes an increasingly interesting research topic. The discovered opinions are useful to many applications [1]. Besides, the opinion analysis technique helps to promote many researches such as automatic summarization [2] and question & answer systems [3]. Many researches on opinion analysis have been reported. Most of them focused on the subjective words extraction [4] and opinion analysis at the document [5] or sentence level [6]. In the recent years, deep opinion analysis, including opinion holder identification and opinion target analysis, becomes a hot research topic [7]. Most existing techniques are camped into three approaches. The fist approach mainly utilizes the knowledge from sentiment lexicon [8]. The second approach is based on machine learning which incorporates linguistics features with a machine learning based classifier [9-11]. The third approach combines sentiment knowledge and FrameNet for opinion analysis [12]. However, current solutions suffer from four major problems. Firstly, many sentiment lexicons assign one polarity label to a sentiment

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 993 – 1000, 2008.
© Springer-Verlag Berlin Heidelberg 2008

word while ignore the abundant occurrences of the sentiment words holding different polarities depending on different contexts. Secondly, the collocation-level and sentence-level opinion-related features are not well studied. Thirdly, the size of annotated opinion corpus is not large enough to support supervised machine learning. Finally, opinion analysis study is still in its infancy [7]. The performance and characteristics of different algorithms cannot be fairly evaluated on the same dataset.

In this paper, an opinion analysis system, namely *OA*, is proposed. *OA* comprises of three modules. Preprocessing and Assignment Module (PAM) performs word segmentation, part-of-speech (POS) tagging and statistics-based collocation extraction. In the Knowledge Acquisition Module (KAM), some word-level and collocation-level linguistic clues for opinion analysis are discovered based on the observation on NTCIR opinion training corpus [10]. Learning algorithms are developed to train the features corresponding to these clues. Meanwhile, new opinion words and their context-dependent polarity behaviors are learned. The Sentence Analysis Module (SAM) incorporates the discovered features into a support vector machine (SVM) based classifier for opinionated sentence identification and polarity determination. Evaluations on NTCIR-6 testing dataset show that *OA* achieved 0.780 (lenient evaluation) and 0.581 (strict evaluation) F-value on opinionated sentence identification, and 0.433 (lenient evaluation) and 0.365 (strict evaluation) F-value on polarity determination, respectively. This performance is promising compared with the evaluated systems in NTCIR-6 (Chinese side) [7].

The rest of this paper is organized as follows. Section 2 briefly reviews the related works. Section 3 presents Preprocessing and Assignment Module. Section 4 describes Knowledge Acquisition Module and Section 5 presents Sentence Analysis Module. Section 6 presents the evaluation results and finally, Section 7 concludes.

2 Related Works

Opinion analysis is an active research topic in text mining and knowledge discovery [13,14]. Early works focus on the identification and polarity determination of sentiment words[4,14,15]. In recent years, opinion analysis techniques are proposed to identify document- and sentence-level opinions from news articles[8], product reviews[1], movie reviews[16] and web blogs[17]. These techniques can be categorized into three major approaches. The first approach mainly utilizes the linguistic knowledge on sentiment words and some opinion-related heuristic rules as the clues for opinion analysis. The typical systems following this approach include [4] on English and[8] on Chinese, respectively. The second approach utilizes machine learning based classifiers and sentiment features, such as sentiment words, word bi-grams, syntactic patterns and topic-relevant features for opinion analysis. The popular used classifiers including Naïve Bayes (NB)[5,10], Maximum Entropy (ME)[9] and Support Vector Machine (SVM) [10,11]. Kim and Hovy[12] proposed a method based on semantic role labeling which follows the third approach. This method utilized a ME-based model to label the semantic roles of opinion related frames in a sentence. The opinion holders and topic are determined by mapping semantic roles to them.

3 Preprocessing and Assignment Module

The word segmentation and POS tagging is always an indispensable preprocessing in Chinese text analysis. A Unicode-based adaptive word segmentation and POS-tagging system, developed by [18], is adopted in PAM. A chunker based on lexicalized Hidden Markov Model is then conducted to recognize the base phrases in the text [19]. Furthermore, some functional components are incorporated to recognize named entities in the sentences including person names, place names and organization names.

The collocation extraction system developed by [20] is adopted. For a given headword, it measures the co-occurrence frequency significance, co-occurrence distribution significance, the non-compositionality and substitutability of all of its co-occurred words and identifies the ones with strong significance as collocations.

4 Knowledge Acquisition Module

4.1 Concepts Related to Opinion Expression

Opinions are usually expressed in some conventional patterns, especially in formal news text. A complete opinion typically consists of five components: (1) **opinion holder**, which is the governor of an opinion and normally refers to a person, a state or an organization as well as the corresponding pronouns; (2) **opinion object**, which is the target of the opinion. Most opinion objects are nouns, noun phrases or pronouns; (3) **opinion word**, which reflects the opinion polarity, i.e. positive, neutral or negative; (4) **opinion operator**, which is the verb indicating an opinion event; and (5) **opinion indicator**, which is the word indicating the orientation of an opinion or the orientation tendency of multiple opinions. In the real txt, no all of the above components always appear in each opinionated sentence.

The opinion operators are the verbs which indicating opinion expression. The typical opinion operators including 警告(warning) 強調(emphasize) and指出 (point out). It is noted that some operators bring opinions such as稱讚(*praise*).

Opinion indicators are mainly conjunctions, adverbs and adverbial phrases, including: (1) negation conjunctions, such as但是(*but, however*) and 儘管(*Though*), indicate that the sentiment of the following clause/sentence is different from the preceding one; (2) continual conjunctions, such as 并且, 而且(and) and 特别 (*especially*), indicate that the sentiment of the following clause/sentence is the same as the preceding one; (3) adverbs and adverbial phrases directly indicate the polarity of the opinionated sentence, e.g. 遺憾的是 (*It is regrettable*); and (4) verbs directly indicate the polarity of the opinionated sentence.

Opinion words play a key role in opinion expression. They are generally classified into: (1) c**ontext-free opinion word** (***CFOW***) whose polarity is constant irrespective of context, e.g. 完美(*perfect*) is absolute positive and 恶劣(*bad*) is negative; (2) **context-dependent opinion word** (***CDOW***) whose polarity is determined by their context. For example, 好笑的is positive when it is used in the context of talk shows (meaning *burlesque*); but it is negative when it is used in the context of politics (meaning *absurd)*; and (3) **object-dependent opinion word** (***ODOW***) is the neutral

word carrying different polarities when associated with different opinion objects. For example, 高(*high*) expresses positive sense when collocating with 性能(*performance*) but brings negative sense when collocating with 債務(*debt*). For practical reasons, this kind of words are always processed as CDOWs.

4.2 Discovery Opinion-Related Linguistic Clues

Linguistic observations on NTCIR opinion training corpus are performed in this study to discovery clues for opinion analysis. This annotated corpus consists of 429 news documents in Chinese. In its 7,980 sentences, 4,716 ones are annotated as opinionated, including 2,735 positive ones, 986 negative ones and 995 neutral ones.

The word-level linguistic clues are studied. Firstly, it is observed that the annotated opinion holders having 13,598 matches in 5,871 sentences in which 3,530 sentences (60.1%) are opinionated. Considering that most opinion holders are nouns, named entities and pronouns. Thus, the recognition of named entities and pronouns is a clue for opinionated sentence identification. Secondly, it is observed that the annotated opinion operators having 1,996 matches in 1,858 sentences in which 1,538 sentences (82.7%) are opinionated. It shows that opinion operator is a good feature for identifying opinionated sentence but the coverage is not high. Thirdly, it is observed that in the 7,484 sentences having the opinion words (44,441 matches), 4,685 ones (62.6%) are opinionated. Meanwhile, that only 31 opinionated sentences (0.57%) have no matched opinion words. It shows that the opinion analysis approach based on opinion lexicon can achieve good recall but the precision is not satisfactory. Next, the status of context-dependent opinion words is observed. Among 3,830 types of annotated opinion words, we find that 384 ones (10.0%) have more than one polarity, i.e. CDOWs. They occur for 4,210 times, which occupies 27.1% of the occurrences of all opinion words. It shows that CDOWs are widely existed. Finally, the status of opinion indicators is observed. In this study, 98 opinion indicators (including 37 negation ones, 22 continual ones, 18 adverbs and 21 verbs) are manually compiled. Among the 581 sentences having opinion indicators, 512 ones (88.1%) are opinionated. Obviously, opinion indicator is a clue.

The collocation-level clues are studied. The collocations between opinion holders and opinion operators are firstly observed. Among the 1,580 sentences having collocations between annotated opinion holders and opinion operators, 94.1% of them are opinionated. This observation distinctly shows that collocation between opinion holder and opinion operator is a better feature for opinion analysis compared to use them separately. Secondly, the collocations related to degree adverbs are studied. Here, the degree adverbs refer to the adverbs carrying degree information, such as 大幅/ad (*greatly*) and 強烈/ad (*strongly*). In this study, 46 degree adverbs from (Lu et al.) are used. Among the 421 sentences having collocations between degree adverbs and opinion words, 398 ones (94.5%) are opinionated. Finally, the collocations between opinion objects and opinion words are studied. We manually annotated 11,726 opinion objects of the collocated opinion words in NTCIR training corpus. It is observed that by using collocated opinion objects as the only feature, the polarities of 98.8% CDOWs can be correctly assigned. This result justifies the '*One Sense per Collocation*' hypothesis applicable to the polarity determination of CDOWs.

4.3 Learning of Opinion Knowledge

The NTCIR opinion training corpus is used as annotated data. As for the raw learning text, the text from NTCIR opinion testing dataset (2,250 documents which consists of 27,841 sentences) is used as the first source. Since its size is rather small, documents relevant to the 32 topics specified by NTCIR corpus are collected from Internet as the second source. It consists of 9,800 documents with 231,832 sentences.

Based on several linguistic resources, an opinion lexicon consisting of 6,924 positive words and 9,627 negative words is built. In which, 1,497 positive words and 2,390 negative ones are manually annotated as CFOWs, and the rest are CDOWs.

Considering that the scale of annotated corpus is limited, we extracted the sentences satisfying any two of following five conditions from raw text:

(1) Sentences contain known CFOWs;
(2) Sentences contain more than three known CDOWs;
(3) Sentences contain collocations between degree adverbs and known CDOWs;
(4) Sentences contain collocations between degree adverbs and operators;
(5) Sentences contain known opinion indicator and known CDOWs.

These sentences are regarded as the ones holding strong polarities and with good confidence. The polarity of these sentences is determined by summarizing the polarity of its opinion words like [8]. These sentences are combined with NTCIR training data (labeled as S_0) to construct an larger opinionated sentence set, labeled as S_1.

Two types of collocations are extracted from S_1, i.e. the collocations between adverbs (out of the known opinion indicators and degree adverbs) and known opinion words, and the collocations between adverbs and known opinion operators in S_1. The adverbs with high collocation frequencies are selected for manually verification of new degree adverbs. Totally, 24 new degree adverbs are obtained. The collocations between opinion holders and their collocated verbs in S_1 are also observed. Suppose a verb v occurs in the opinionated sentences for t_{op} times while it occurs in the whole training text for t times, the probability of verb v that serves as opinion operators, labeled as $P_op(v)$, is estimated by,

$$P_op(v) = t_{op} / t \tag{1}$$

The verbs with top-values of $P_op(v)$ and occurrences greater than a empirical threshold are selected as new opinion operators.

The context behaviors of known opinion words are learned from S_1. For each opinion word *ow*, suppose it occurs in S_1 for t_{ow} times, in which it holds positive polarities for $t_{_pos}$ times, neutral polarities for t_{neu} times and negative polarities for t_{neg} times, the probabilities that it is determined as positive, neutral and negative, named as *polarity scores*, with the label of $P_pos(ow)$, $P_neu(ow)$, and $P_neg(ow)$, respectively. In which, $P_pos(ow)$ is calculated by,

$$P_pos(ow) = t_{pos} / t_{ow} \tag{2}$$

$P_neu(ow)$, and $P_neg(ow)$ are calculated in the same way. Values of these scores range from 0 to 1 and larger value indicates stronger polarity. Furthermore, the nouns collocated with *ow* in the sentences of S_1 are collected. They are divided to three sets

corresponding to different polarities hold by *ow*. The lists of collocated nouns are in turn record in the opinion knowledge database (associate with record of *ow*).

5 Sentence Analysis Module

Describing the sentences as a set of values of opinion-related features, these sentences are transferred to points in a multi-dimension space. The SVM classifier attempts to find out hyperplane in the feature space which separates the positive training examples from the negatives ones. The three groups of features are designed or selected, which is given in Table 1. Using S_1 as training data, the trained classifier analyzes each input sentence and determines its polarity as the output.

Table 1. List of Features for Opinion Analysis

Group1. The features related to punctuations
The presence of direct quote punctuation "「"and "」"
Group 2: The word-level and entity-level features
The summary of the probabilities of known opinion operators. Following Equation 1
The sentence polarity score. Following Equation 4 and values [0,1]
The presence of a person name or a person entity.
The presence of a named entity of location, country and organization.
The presence of pronoun.
The presence of the nouns appearing in the title of the document.
The presence of the named entity appearing in the title of the document.
The presence of known opinion indicators.
The presence of known degree adverbs.
The percentage of entities and pronouns in the sentence, values [0, 1]
Group 3: The collocation-level features
The presence of collocations between person names/entity and opinion operators.
The presence of collocations between pronouns and opinion operators.
The presence of collocations between nouns or named entities and opinion words.
The presence of collocations between pronouns and opinion words.
The presence of collocations between degree adverbs and opinion words.
The presence of collocations between degree adverbs and opinion operators.

6 Evaluations

The NTCIR-6 opinion testing dataset with the gold answers is adopted for evaluation. It consists of 27,841 sentences in 2,250 documents corresponding to 28 topics. In the 17,269 opinionated sentences, 6,894 ones are positive, 5,889 ones are neutral and 4,486 ones are negative. Besides *OA*, we evaluated two typical existing algorithms. The first one, labeled as *A1*, is similar to NTU system [8] which estimates the polarity of a sentence by using the summary of its word polarities. The second algorithm, labeled as *A2*, is based on a SVM classifier which regards the occurrences of the words in opinionated and non-opinionated sentences, respectively, are features.

The performances on opinionated sentence identification are firstly evaluated. The achieved performances are listed in Table 2. In the following tables, *P* is for precision, *R* is for recall and *F* is for F-value.

Table 2. Performances on Opinionated Sentence Identification

System	Lenient Evaluation			Strict Evaluation		
	P	*R*	*F*	*P*	*R*	*F*
A1	0.654	**0.852**	0.740	0.237	**0.920**	0.377
A2	0.741	0.532	0.619	0.453	0.384	0.416
OA	**0.851**	0.721	**0.780**	**0.497**	0.698	**0.581**

It is observed *A1* achieves good recall. However, its precision is difficult to improve. *A2* does not achieve a high performance. It is because that this algorithm discards many opinion-related linguistic features. By incorporating such kinds of features, its performance will improve. *OA* outperforms these two algorithms, which attributes to the use of effective opinion-related features and learning capability.

The performance on sentence polarity determination achieved by the same systems with the same settings are evaluated. The achieved performances are listed in Table 3.

Table 3. Performances on Polarity Determination

System	Lenient Evaluation			Strict Evaluation		
	P	*R*	*F*	*P*	*R*	*F*
A1	0.346	**0.441**	0.388	0.119	0.649	0.201
A2	0.284	0.403	0.333	0.095	0.461	0.157
OA	**0.596**	0.341	**0.433**	**0.253**	**0.659**	**0.365**

It is shown that for *A1*, an expanded lexicon does not ensure the performance improvement on polarity determination because a larger lexicon but without sufficient and accurate statistics sometime affects the system. In these systems, *OA* performs best both on lenient evaluation and strict evaluation.

7 Conclusions

In this paper, learning techniques are proposed to discover features for opinion analysis from both annotated and raw text. These features are incorporated into a SVM-based classifier for opinion analysis. Evaluations on NTCIR-6 opinion testing dataset show that our system achieves good overall performance. This result justifies that the learning of opinion-related linguistic features, especially context-dependent opinion words and the collocation-level features improving opinion analysis.

Acknowledgments. This research is supported by Hong Kong Polytechnic University (A-P203), a CERG Grant (5087/01E) and The Chinese University of Hong Kong under the Direct Grant Scheme project (2050417) and Strategic Grant Scheme project (4410001).

References

1. Hu, M.Q., Liu, B.: Mining and Summarizing Customer Reviews. In: ACM SIGKDD 2004, Seattle, pp. 168–177 (2004)
2. Hu, M.Q., Liu, B.: Opinion Extraction and Summarization on the Web. In: NCAI 2006, Boston (2006)
3. Yu, H., Hatzivassiloglou, V.: Towards Answering Opinion Question: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences. In: EMNLP 2003 (2003)
4. Hatzivassiloglou, V., McKeown, K.: Predicting the Semantic Orientation of Adjectives. In: ACL1997, Madrid, pp. 174–181 (1997)
5. Pang, B., Lee, L.L., Vaithyanathan, S.: Thumbs up? Sentiment Classification Using Machine Learning,Techniques. In: EMNLP 2002, Philadelphia, pp. 79–86 (2002)
6. Riloff, E., Wiebe, J., Wilson, T.: Learning Subjective Nouns Using Extraction Pattern Bootstrapping. In: CoNLL 2003, pp. 25–32 (2003)
7. Seki, Y., Evans, D.K., Ku, L.W., Chen, H.H., Kando, N., Lin, C.Y.: Overview of Opinion Analysis Pilot Task at NTCIR-6. In: NTCIR-6, Japan, pp. 456–463 (2007)
8. Ku, L.W., Wu, T.H., Lee, L.Y., Chen, H.H.: Construction of an Evaluation Corpus for Opinion Extraction. In: NTCIR-5, Japan, pp. 513–520 (2005)
9. Pang, B., Lee, L.L.: A Sentiment Education: Sentiment Analysis Using Subjectivity Summariza-tion based on Minimum Cuts. In: ACL 2004, Spain, pp. 271–278 (2004)
10. Lin, W.H., Wilson, T., Wiebe, J., Hauptmann, A.: Which Side are You on? Identifying Perspectives at the Document and Sentence Level. In: CoNLL 2006, pp. 109–116 (2006)
11. Riloff, E., Patwardhan, S., Wiebe, J.: Feature Subsumption for Opinion Analysis. In: EMNLP 2006 (2006)
12. Kim, S.M., Hovy, E.: Extracting Opinions Expressed in Online News Media Text with Opinion Holders and Topics. In: COLING-ACL 2006 (2006)
13. Xia, Y.Q., Xu, R.F., Wong, K.F., Zheng, F.: The Unified Collocation Framework for Opinion Mining. In: IEEE ICMLC 2007, Hong Kong, pp. 844–850 (2007)
14. Turney, P., Littman, M.: Measuring Praise and Criticism: Inference of Semantic Orientation from Association. ACM Trans. Information Systems 4(21), 315–346 (2003)
15. Kamps, J., Marx, M., Mokken, R.J., Rijke, M.: Using WordNet to Measure Semantic Orientation of Adjectives. In: LREC 2004 (2004)
16. Chaovalit, P., Zhou, L.: Movie Review Mining: a Comparison between Supervised and Unsupervised Classification Approaches. In: HICSS 2005 (2005)
17. Zhang, W., Liu, S., Meng, W.Y.: Opinion Retrieval from Blogs. In: ACM CIKM 2007, Lisbon, pp. 831–840 (2007)
18. Lu, Q., Chan, S.T., Xu, R.F.: A Uni-code based Adaptive Segmentor. In: Dignum, F.P.M. (ed.) ACL 2003. LNCS (LNAI), vol. 2922, pp. 164–167. Springer, Heidelberg (2004)
19. Fu, G.H., Xu, R.F., Luke, K.K., Lu, Q.: Chinese Text Chunking Using Lexicalized HMMs. In: Yeung, D.S., Liu, Z.-Q., Wang, X.-Z., Yan, H. (eds.) ICMLC 2005. LNCS (LNAI), vol. 3930, pp. 7–12. Springer, Heidelberg (2006)
20. Xu, R.F., Lu, Q., Wong, K.F., Li, W.J.: Classification-based Chinese collocation extraction. In: IEEE NLPKE 2007, Beijing, pp. 308–315 (2007)

An Indexing Matrix Based Retrieval Model

Xu Zhao and Zongli Jiang

The Laboratory of Computer Software and Theory
Beijing University of Technology, PingLeYuan 100,
ChaoYang District, PR China
zhaoxu166@gmail.com, jiangzl@bjut.edu.cn

Abstract. The Latent Semantic Indexing based standard retrieval usually computes semantic similarity between the query and every document, which is sensitive to the noise. Furthermore, it will take much more time when the scale of the corpus becomes larger. We propose a retrieval method which is based on an indexing matrix. The noises can be wiped off at a certain extent and it can get a better result than the traditional ways. The time cost of our method is much less than the standard retrieval method.

Keywords: Indexing Matrix, Term Similarity, Latent Semantic Indexing, Document Representation.

1 Introduction

Along with the rapid development of the INTERNET, information lying in the web increases at a progressional speed. Facing these vast and complicated resources, people are expecting a more efficient retrieval method.

Traditional terms based Vector Space Model (VSM) uses n terms to form a document vector $D_i = \{d_{i1}, d_{i2}, \ldots d_{in}\}$, where d_{ij} (j=1,2…n) represents the weight of the j-th term [1]. By transforming the unstructured raw documents into some vectors, some mathematical disposals can be done easily. This expression is undemanding and easy to compute. It considers that two documents' similarity is based on how many terms belong to both of them and computes two documents' similarity according to their term frequency hence neglects the more important semantic relationship in the nature language. Meanwhile, VSM is based on an assumption that different terms are unrelated with each other. But in reality, there are more or less some semantic relationships between different words, so the assumption is difficult to satisfy.

Whereas VSM's disadvantage, Latent Semantic Indexing (LSI) has been developed [2]. LSI believes that there is an underlying semantic structure between the words lying in the corpus, and the semantic structure is veiled by the context. By the statistical analysis, LSI can capture this latent semantic relationship; consequently make the retrieval results better.

2 Latent Semantic Indexing and Standard Retrieval Method in the Semantic Space

As a document, all its words have some semantic relationships between each other more or less. A source of information about semantic similarity in the given context is

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1001 – 1008, 2008.
© Springer-Verlag Berlin Heidelberg 2008

co-occurrence analysis: if two terms co-occur in documents very frequently (in a given corpus) they can be considered as semantically related. Incorporating co-occurrence information in a learning system for exploiting semantic similarity between words would seem to be a very expensive task. However, a technique has been developed in the information retrieval literature that can extract this information automatically [3]. It is Latent Semantic Indexing (LSI).

LSI projects all documents into a space with "Latent Semantic Dimensions", where co-occurring terms are projected in similar directions, while non co-occurring ones are projected in very different directions. In such a space two documents can have a high similarity even if they do not share the same terms(as long as they each use terms frequently co-occurring in the corpus) [3].Therefore, the key point is to depict the relationships between the terms and the documents. Usually we can use document-term matrix [4]. For m documents and n terms, let

$$C = \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \cdots & \cdots & \cdots \\ C_{m1} & \cdots & C_{mn} \end{bmatrix}$$

where rows are indexed by the documents of the corpus and columns by the terms. The (j,i)-th entry of C gives the frequency of term t_i in document d_j. By weighting the terms, we can get a more reasonable relationship between terms and documents. It is obvious that different weighting methods can lead to different impacts [5], [6].

We transpose C to get the term-document matrix C^T. Then we apply Singular value decomposition (SVD) [7] to C^T and get

$$C^T = UQV^T, \tag{1}$$

where $U^TU=V^TV=I_n$ and $Q = diag(\sigma_1,...,\sigma_n), \sigma_i > 0$ for $1 \leqslant i \leqslant r$ (r is the rank of matrix C^T), $\sigma_j = 0$ for $j \geqslant r+1$. The first r columns of the orthogonal matrixes U and V define the orthonormal eigenvectors associated with the first r nonzero eigenvalues of C^TC and CC^T, respectively. Now we need to choose a value k ($0 < k \leqslant r$) to approximate C^T by $C_k{}^T=U_kQ_kV_k{}^T$, and the space spanned by U_k is the semantic space we want to get. Notice that C^TC is the covariance matrix of the corpus, so the first k diagonal elements of matrix Q are the variances corresponding to the directions of the principal axes that span U_k. So, the larger the dimension k of the subspace U_k, the greater percentage of the variance that is captured.

According to this idea, we choose k ($0 < k \leqslant r$) that makes $\sigma_k >> \sigma_{k+1}$, which means that we want to capture the variances as many as the original dataset, and we get

$$C_k{}^T = U_kQ_kV_k{}^T \qquad (0 < k \leqslant r), \tag{2}$$

where U_k is made up of the first k columns of U ,and Q_k is made up of the first k rows and the first k columns of Q, and V_k is made up of the first k columns of V. So, $C_k{}^T$ is the approximate representation of C^T in the least squares sense. We can project all the documents from the original space into the semantic space spanned by U_k, and get

every document's representation in the semantic space. In this representation, we notice that the covariance matrix is diagonal:

$$\frac{1}{n} U_k^T C^T C U_k = \frac{1}{n} U_k^T U Q U^T U_k = \frac{1}{n} Q_k . \quad (3)$$

So, the latent semantic dimensions are approximately unrelated with each other.

Similarly, we view a term as a document that contains only this term, and project this term into the semantic space after the same weighting. Now all the documents and the terms are in the same space. In this semantic space, semantically correlative terms or terms occur in correlative documents will be close to each other, even if they never co-occur in a document. This means that even though some documents do not contain the query, they will appear at the adjacent place in the semantic space.

When querying, we represent the query as a vector in the original space, and apply the same weighting method, then project it into the semantic space. By computing its similarity to every document existing in the semantic space and ranking the documents by their similarity to the query, the documents whose similarity to the query exceed a threshold are returned to the user.

The standard retrieval method is easy and intelligible. But as a document, it is unlikely that all its terms are equally important to it. Furthermore, the terms which are appeared in only one document may be noise. Hence, the document may deflect from their topic in the semantic space after projecting, which makes the retrieval somewhat inaccurate to a query. Also the standard retrieval method needs to compute the similarity between the query and all the documents, which will cost too much time.

3 Indexing Matrix Based Retrieval

To deal with the problem presented above, we propose an indexing matrix based method. We name it IMBR (Indexing Matrix Based Retrieval). Its basic idea is to identify the documents corresponding to a query by the semantic similarity between terms. First, we centralize the term-document matrix to locate the origin at the centroid of the corpus. We just the same use C^T to represent the term-document matrix which has been centralized. By applying SVD to C^T, we can get the derived k ranked approximate matrix $C_k^T = U_k Q_k V_k^T$. Then, we project all terms from the original space into the semantic space and compute the semantic similarity between every two terms in the semantic space, and preserve the results in a matrix named SI. Let the document be represented by the centroid of all its topic terms, namely the mean value vector of all its topic terms, and get the indexing matrix. When a query is coming, we just need to visit the indexing matrix, and get the documents corresponding to the query terms. Then we sort the documents according to a synthetical index that indicates the most reasonable relationship between the documents and the terms, and return to user.

3.1 How to Construct the Indexing Matrix

Assume that term I and J are represented as I=(0,0,…,i,0,…0) and J=(0,0,…,j,0,…0), respectively, when I≠J, i and j are at the different dimension. The only non-zero

dimension indicates the weight of the term to this "document". In this way, we have the representation of term I and J in the semantic space:

$$I_{semantic} = I * U_k, \tag{4}$$

$$J_{semantic} = J * U_k. \tag{5}$$

where $I_{semantic}$ and $J_{semantic}$ are all k ranked row vectors. Let $H = U_k * U_k^T$, the inner product of $I_{semantic}$ and $J_{semantic}$ can be computed as: $(I_{semantic}, J_{semantic}) = I * U_k * U_k^T * J^T = I * H * J^T = i * j * H_{ij}$. Consequently, the similarity between I and J in the semantic space can be computed as follows:

$$\begin{aligned}
&\mathrm{COS}(I_{semantic}, J_{semantic}) \\
&= (I_{semantic}, J_{semantic}) / \sqrt{(I_{semantic}, I_{semantic})(J_{semantic}, J_{semantic})} \\
&= i * j * H_{ij} / \sqrt{i * i * H_{ii} * j * j * H_{jj}} \\
&= H_{ij} / \sqrt{H_{ii} * H_{jj}}.
\end{aligned}$$

Let

$$SI_{ij} = H_{ij} / \sqrt{H_{ii} * H_{jj}}. \tag{6}$$

We obtain the matrix SI whose elements are the similarity between terms in the semantic space.

Every row of SI can be regarded as a vector which represents a term. Every entry of the vector is the similarity between this term and the other terms.

What should be done next is to represent every document by its similarity with every term. Notice that not all terms belong to a document are positive to this document, some terms may be noise.

Here, in view of the computation efficiency we deem one term which appears more than twice in a document is the term that has a topically indicative effect, and name it "topic term". (Of course a term that appears more than 3 times or more times in a document also can be a criterion for selecting "topic term", but it will take a bit of time when compute the indexing matrix). Meanwhile, we view the terms which appear only one time in a document as noise. In this way, we can represent a document as the centroid of its topic terms. Assume that document d contains h topic terms t_1、t_2、…、t_h which appear $f(t_i)$ times in d, respectively. Then d can be represented as

$$doc = \sum_{i=1}^{h} (f(t_i) * SI(t_i)) / \sum_{i=1}^{h} f(t_i), \tag{7}$$

where SI (t_i) corresponds to the row vector of term t_i in the matrix SI. Vector doc is the centroid of all d's topic terms. By this representation, every entry of doc is a similarity between document d and a term. Obviously, the entries corresponding to d's topic terms will be larger than the noise terms.

According to this way, we compute all m document vectors, and compose m*n matrix B. Every row of B represents a document, and the (i,j)-th entry of B gives similarity between the i-th document and the j-th term. Let

$$A=B^T . \tag{8}$$

Obviously, the (i,j)-th entry of A gives the similarity between i-th term and the j-th document, and the i-th row of A gives the similarity between the i-th term and all the documents in the corpus. By sorting every row of A, respectively, we can get the order of every term with all the documents.

3.2 The Algorithm

Based on the idea stated in section 3.1, the algorithm can be concluded as follows:

1 Pretreat the corpus, and get the term-document matrix C^T;
2.Centralize C^T, and apply SVD to get the k ranked approximated matrix $C_k^T=U_kQ_kV_k^T$;
3.Use equality (6) to obtain matrix SI;
4.Compute every document's centroid vector of all its topic terms, and store the result in the matrix B;
5. Let $A=B^T$, and sort every row of A to get the indexing matrix.

When a new query is coming, the only thing need to be done is to find the rows corresponding to the query in the indexing matrix A, and intersect the results. When intersecting, we must consider the problem of sorting. In our experiment, we locate a document in the result list by mean scores.

4 Experiments and Discussion

The experiments are conducted on a subset of the TanCorp-12 collection [8], [9]. There are totally 3813 documents in our subset. We use Latent Semantic Indexing based standard similarity retrieval (Standard LSI Retrieval) as a benchmark.

The results indicate that comparing to Standard LSI Retrieval, IMBR iwis improves the retrieval performance to some extent. We show two representative experiments here to illustrate the effect. Experiment 1 compares the accuracy and the time cost of the retrieval between Standard LSI Retrieval and IMBR; experiment 2 compares the ability of resisting the noises.

Experiment 1 first use expected cross entropy to conduct feature selection. 7024 terms have been selected, and all the documents in the corpus are represented by vectors using these selected terms. Then TEF-WA[10] is applied to weight the terms and get the term-document matrix. After this step, we centralize the term-document matrix to locate the origin at the centroid of the corpus. By applying SVD to the term-document matrix, we derive the latent semantic space. In this semantic space, Standard LSI Retrieval can be conducted. Also, using the algorithm given in section 3 to process the corpus, we can derive the indexing matrix A, and use A to conduct IMBR.

Concretely we choose some terms to form a query. Then Standard LSI Retrieval and IMBR are applied to this query, respectively. Here, we deem the first 20 results are the ones which are viewed as the best results by the two methods, and they can

illuminate the retrieval quality of the two methods best. Hence, we only compare these first 20 results here.

In query 1 we choose “健康营养”(Means health and nutrition) as the query. The retrieval results are showed in table 1.

Table 1. Using “健康营养” as the query

	Standard LSI Retrieval	IMBR
The position of the first directly related document	2	1
total related documents	18	20
directly related documents	8	14
semantically related documents	10	6
irrelated documents	2	0
overlapped documents	6	
	Directly related	6
	Semantically related	0
Time cost(ms)	2063	36

In query 2 we choose “姚明篮球”(Means MingYao and basketball) as the query. The retrieval results are showed in table 2.

Table 2. Using “姚明篮球” as the query

	Standard LSI Retrieval	IMBR
The position of the first directly related document	5	3
total related documents	18	20
directly related documents	5	13
semantically related documents	13	7
irrelated documents	2	0
overlapped documents	5	
	Directly related	3
	Semantically related	2
Time cost(ms)	2379	42

From the result we can see that comparing to Standard LSI Retrieval, IMBR improve the performance at a certain extent. From the results of query 1, we can see that all the 20 documents found by IMBR are related documents, but there are 2 documents found by Standard LSI Retrieval are irrelated documents. Further, notice that there are 14 directly related documents found by IMBR —almost twice as many as the number of directly related documents found by Standard LSI Retrieval. Meanwhile, 6 documents are overlapped documents, and all these documents are directly related documents, which mean that nearly all the directly related documents found by Standard LSI Retrieval are found by IMBR.

Also the time cost of Standard LSI Retrieval is 2063ms, but IBMR only need 36ms to do the same job. This is because Standard LSI Retrieval needs to compute the similarity between the query and every document in the corpus, whereas IMBR only needs to intersect several rows of the indexing matrix. When the scale of the corpus becomes larger, the difference of the time cost between Standard LSI Retrieval and IMBR will become more clearly, since the time cost of Standard LSI Retrieval is proportionable to the scale, but IMBR is not the same. Generally speaking, the number of the related documents to a term has an upper limit. When retrieving, we can choose a threshold according to the scale of the corpus to control the number of the documents related to every term. So the time cost needed by IMBR increases when the number of related documents becoming larger, not the scale of the corpus. Therefore, the time spending of IMBR may not increase sharply when the scale of the corpus becomes larger.

Similar results can be seen in query 2 (Table 2). The number of directly related documents of IMBR is almost 3 times as many as Standard LSI Retrieval. Also the time cost of IMBR is much less than Standard LSI Retrieval.

Synthesize the above results we can see that IMBR exceeds Standard LSI Retrieval by comparing the relative extent of the retrieval results; this is embodied by the number of directly related documents and the time cost, which indicates that IMBR certainly improves the performance of retrieval.

In experiment 2, we choose “Beckham” and “advertisement” to form a query. By an auxiliary experiment, we find that there are only 5 documents which contains both of the two terms in our subset, and among these 5 documents there is only one document that is directly related to Beckham’s advertisement, and the term “Beckham” and “advertisement” just co-occur in the other 4 documents. By Standard LSI Retrieval, the first top 10 results are all irrelated, but IMBR digs the sole related one out successfully and locates it at the first place of the result list. This implies that IMBR can resist the effect of the “noise term” much more potently than Standard LSI Retrieval at a certain extent.

5 Conclusions

LSI based standard retrieval method needs to project the query and all the documents into the semantic space, and compute the similarity between the query vector and all the documents vectors. But due to the existence of the “noise term”, retrieval results may be disturbed in some degree. Meanwhile, its time spending is pro rata to the scale of the corpus. A new indexing matrix based retrieval method IMBR is proposed. IMBR overcomes the effect of the “noise term” to some extent, meanwhile, the time cost of our method is much less than the standard retrieval method.

References

1. Salton, G., Wong, A., Yang, C.S.: A Vector Space Model for Automatic Indexing. Communications of the ACM 18(11), 613–620 (1975)
2. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by Latent Semantic Analysis. J. Amer. Soc. Inf. Sci. 1(6), 391–407 (1990)

3. Cristianini, N., Lodhi, H., Shawe-Taylor, J.: Latent Semantic Kernels for Feature Selection. NeuroCOLT Working Group (2000), http://www.neurocolt.org
4. Shawe-Taylor, J., Cristianini, N.: Kernel Methods for Pattern Analysis. Cambridge University Press, Cambridge (2004)
5. Dumais, S.T.: Improving the Retrieval of Information from External Sources. Behav. Res. Meth. Instr. Comput. 23, 229–236 (1991)
6. Debole, F., Sebastiani, F.: Supervised Term Weighting for Automated Text Categorization. In: SAC 2003, pp. 784–788. ACM Press, New York (2004)
7. Berry, M.W., Dumais, S.T., O'Brien, G.W.: Using Linear Algebra for Intelligent Information Eetrieval. SIAM: Review 37, 573–595 (1995)
8. Tan, S., et al.: A Novel Refinement Approach for Text Categorization. ACM CIKM (2005)
9. Tan, S., Cheng, X., Ghanem, M., Wang, B., Xu, H.: A Novel Refinement Approach for Text Categorization. CIKM, 469–476 (2005)
10. Tang, H.L., Sun, J.T., Lu, Y.C.: The Weight Regulation Tchnology of Set Evaluation Function in Text Categorization. Journal of Computer Research and Development 42(1), 47–53 (2005)

Real-Time Intelligent End-Diastolic and End-Systolic Image Retrieval from Left Ventricular Angiograms Using Probability Density Propagation and Data Fusion

Wei Qu and Yuan-Yuan Jia

Siemens Medical Solutions USA Inc.
2501 North Barrington Road, Hoffman Estates, IL, USA 60192
weiqu@siemens.com, jiayuanusa@gmail.com

Abstract. Left ventricular analysis is an important cardiac examination, which is widely used in the cath lab. However, due to the intrinsic difficulties of X-ray angiographic images, manual inspection is still needed in the current clinical work flow to retrieve an end-diastolic (ED) and an end-systolic (ES) image from hundreds of frames. Obviously, this process is very tedious and time consuming. This paper presents a novel data fusion based intelligent approach for ED/ES image retrieval. By exploiting the latest image analysis theory and signal processing techniques, it achieves promising performance on the real clinical data.

Keywords: Image Retrieval, Data Fusion, Medical Image Processing.

1 Introduction

Left ventricular angiography has been widely used to evaluate clinical cardiac functions such as ejection fraction, stroke volume, and wall motion abnormalities for many years [1]. During this diagnosis, X-ray opaque contrast dye is injected into a patient's left ventricle in order to visualize its variation. Medical imaging system is used to capture an X-ray image sequence covering 5 to 10 cardiac cycles. In order to analyze the clinical parameters, an end-diastolic (ED) image where the left ventricle is fully filled and an end-systolic (ES) image where the ventricle is maximally contracted have to be retrieved from the sequence. After that, further image analysis such as contour segmentation may be done to determine the ventricular volume [2]. Since ED/ES image retrieval is the first step in this procedure, the selection accuracy is very important to all later analysis. In the current clinical work flow, ED and ES images are manually selected, which is not only time consuming but also sensitive to different persons at different time. Intelligent ED and ES image retrieval has the potential to greatly save users' time, reduce their burden, and increase the analysis accuracy. However, due to the intrinsic difficulties in X-ray left ventricular images such as low contrast, noisy background, and dramatically changing left ventricular shape, it is a very challenging task and still remains unsolved. In [3], Qu et al. proposed a method

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1009–1016, 2008.
© Springer-Verlag Berlin Heidelberg 2008

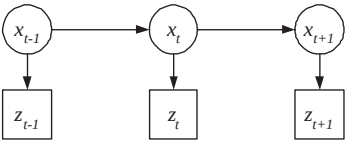


Fig. 1. The Hidden Markov Model used for left ventricular angiograms

called AutoEDES for automatic ED and ES frame selection in angiograms. To the best of our knowledge, it's the first paper investigating this issue using only image data. Although very effective in detecting ED/ES images on most normal testing data, we found that it doesn't work well for the data with poor image quality. Moreover, extra-systolic phenomenon is not considered.

In this paper, we extend the previous work [3] and propose a novel data fusion based intelligent framework for real-time ED and ES image retrieval from left ventricular angiograms. It applies a much deeper analysis of image data and exploits both prior domain knowledge and ECG signal achieving a more robust performance. The rest of the paper is organized as follows: image data analysis, ECG signal processing, and domain knowledge integration are discussed in Sections 2, 3, 4, respectively; Section 5 presents our data fusion framework; Section 6 gives the experimental results; Finally, we conclude the paper.

2 Image Analysis Using Bayesian Density Propagation

The left ventricular angiogram sequence can be modeled by a Hidden Markov Model (HMM) [4] as illustrated in Fig. 1. Circle nodes represent the states which are defined as the area value of left ventricle. t is the time index. Square nodes represent the image observations. The directed link between two consecutive states represent the dynamics $p(x_t|x_{t-1})$. It is a Markov chain where each state only conditionally depends on the previous state. The directed link between the state and its observation is the local likelihood density $p(z_t|x_t)$. Compared to the graphical model in [3], we adopt a simpler formulation in this paper in order to avoid complicated background modeling and thus make the framework more generic. Based on the HMM, it is easy to derive the Bayesian propagation rule,

$$p(x_{1:t}|z_{1:t}) = \frac{p(x_t, z_t|x_{1:t-1}, z_{1:t-1})}{p(z_{1:t})} p(x_{1:t-1}, z_{1:t-1}) \tag{1}$$

$$= \frac{p(x_t, z_t|x_{t-1})}{p(z_t)} p(x_{1:t-1}|z_{1:t-1}) \tag{2}$$

$$= \frac{p(z_t|x_t, x_{t-1})p(x_t|x_{t-1})}{p(z_t)} p(x_{1:t-1}|z_{1:t-1}) \tag{3}$$

$$\propto p(z_t|x_t)p(x_t|x_{t-1})p(x_{1:t-1}|z_{1:t-1}). \tag{4}$$

In (2) and (4), Markov properties have been applied. By this update equation, the posterior density of state x_t can be recursively estimated. Since the left ventricle

changes its shape continuously and smoothly between the consecutive frames, we can thus model the dynamics $p(x_t|x_{t-1})$ by a Gaussian distribution $\mathcal{N}(0, \sigma_d^2)$. In (4), how to model the observation likelihood $p(z_t|x_t)$ is not trivial because the exact observation z_t is unknown. Effective feature extraction techniques are needed to estimate the observation information out of the image resources. We have designed two efficient observation likelihood models as follows.

2.1 Histogram-Based Observation Likelihood Model

In the beginning of a left ventricular angiogram, there is no dye and the background doesn't change dramatically. This information can be exploited to extract the background information. We propose a histogram-based likelihood model to against the background noise. Firstly, a number of images, say, n_{BG}, are selected from the beginning of the sequence. Secondly, in each frame, we select a predefined region of interest (ROI) with size of w_h by h_h and build up the histogram based on this region. After that, a histogram template $\hbar_0$ is calculated by averaging the n_{BG} histograms. This histogram represents the average pixel intensity distribution of the background without any dye. Finally, we calculate the local observation likelihood as follows:

$$p(z_t^h|x_t) = \frac{1}{k_h} exp\{-\frac{(x_t - \int_i sign(\hbar_t^i - \hbar_0^i))^2}{2\sigma_h^2}\} \tag{5}$$

where

$$sign(\cdot) = \begin{cases} 1 & \text{if } (\hbar_t^i - \hbar_0^i) > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

i is the bin index, σ_h is the variance, and k_h is a normalization factor.

2.2 Scene-Based Observation Likelihood Model

The histogram-based observation likelihood has the advantage to capture the global pixel intensity variation. Moreover, since the histogram is independent on the pixels' spatial information, this model has good tolerance to background noise. However, we found that the histogram-based likelihood model suffers from global lightness variation and has difficulty to accurately capture the small region change especially during the cardiac ES phase. This motivates us to propose another scene-based observation likelihood model. Specifically, a number of n'_{BG} images in the beginning of the whole angiogram are chosen. This number can be same or different as n_{BG}. Then a much smaller window with size of w_s by h_s instead of the previous ROI, where $w_s < w_h$, and $h_s < h_h$, is automatically selected by the algorithm for all the n'_{BG} images in order to save the computational cost. This window should cover the most possible region where the left ventricle can be. For each pixel in the window, a background template r_0 is generated by selecting the minimal intensity value among the n'_{BG} candidates. Intuitively, this template reflects the blackest situation of each pixel during the beginning of the left ventricular angiogram. After that, we do background subtraction within

the image window for each frame, $r_t - r_0$, where r_t is the image window at time t. Then we calculate the scene-based observation likelihood as follows,

$$p(z_t^s|x_t) = \frac{1}{k_s} exp\{-\frac{(x_t - \int_j sign(r_t^j - r_0^j))^2}{2\sigma_s^2}\} \tag{7}$$

where

$$sign(\cdot) = \begin{cases} 1 & \text{if } (r_t^j - r_0^j) > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

j is the pixel index, σ_s^2 is the variance constant, k_s is a normalization factor.

2.3 Joint Posterior Density

The two observation likelihood models can be integrated together by calculating the joint posterior density,

$$p(x_{1:t}|z_{1:t}^h, z_{1:t}^s) \propto p(z_t^h, z_t^s|x_t)p(x_t|x_{t-1})p(x_{1:t-1}|z_{1:t-1}^h, z_{1:t-1}^s) \tag{9}$$

$$= p(z_t^h|x_t)p(z_t^s|x_t)p(x_t|x_{t-1})p(x_{1:t-1}|z_{1:t-1}^h, z_{1:t-1}^s) \tag{10}$$

where we assume that the observation likelihoods are conditionally independent.

3 ECG QRS-Complex Detection

During the left ventricular angiography, ECG signal is usually recorded simultaneously. Compared with the image data, ECG signal is independent to the image quality and the dye contrast. Since the R-wave of an ECG signal corresponds to the ED phase in one cardiac cycle, and the T-wave corresponds to the ES phase, ECG signal has potential to be used for the ED/ES selection. Unfortunately, due to the inherent characteristics of the ECG signal, although many ECG-based ED/ES methods are available in the literature [5], several limitations prevents the direct use of these ECG QRS-complex detection algorithms in the ED/ES image retrieval problem. Specifically, (1) ECG signal can not show when the dye is injected into left ventricle. Thus, it could not find the desirable cardiac cycle with good dye contrast; (2) ECG QRS-complex detection methods may generate both false positives and false negatives; (3) ECG signal may have synchronization errors; (4) T-wave is difficult to be detected accurately because it is easy to be corrupted by the noise.

We use ECG signal to generate ED candidates by R-wave detection [5]. Based on these ED candidates, the ED distribution density of ECG signal can be defined as a Gaussian mixture,

$$f(t) = \sum_l \mathcal{N}(t_l, \sigma_e^2) \tag{11}$$

where l is the index of ED candidates, σ_e^2 is a variance constant.

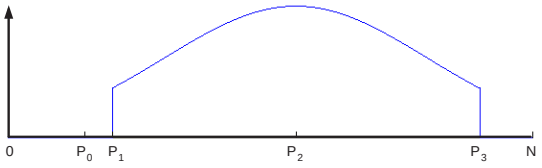


Fig. 2. The density distribution based on prior domain knowledge

4 Information Integration of Prior Domain Knowledge

Prior domain knowledge has to be considered during the retrieval of optimal ED/ES images. For example, extra-systolic phenomenon usually occurs after dye is injected into the left ventricle, where the cardiac cycle in this period has an additional ES phase, and the ED phase is also 20% to 25% larger than usual. Such kind of ED and ES images should not be retrieved though they may have very good dye contrast. How to combine these prior domain knowledge into the ED/ES image retrieval framework is not trivial. In this paper, we propose a novel way by integrating all domain information into a probability density. Fig.2 shows the main idea, where N is the total frame number of the left ventricular angiogram. P_0 is the frame point when dye is injected into left ventricle. The cardiac cycle around P_0 is the place where the extra-systolic phenomenon may begin to occur. In order to avoid any selection inside extra-systolic cardiac cycle, we use point P_1 to represent the starting point of ED/ES image retrieval search range. The distance $|P_0P_1|$ is a predefined constant. P_3 is a cutoff point to avoid any ED/ES image selection from the end of the sequence, where the dye fades away and the left ventricle area doesn't have enough contrast anymore. The distance $|P_3N|$ is a empirical number proportional to the total frame number N, and the sequence frame rate f_r. P_2 is the middle point between P_1 and P_3. The prior density of domain knowledge thus can be defined as

$$p_0(t) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_p} exp\{-\frac{(t-P_2)^2}{2\sigma_p^2}\} & \text{if } P_1 \le t \le P_3; \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

where σ_p^2 is a variance constant.

5 Data Fusion Based ED/ES Image Retrieval

In this section, we present our data fusion method for ED and ES image retrieval, respectively.

Without loss of generality, we can assume that the image data, ECG signal and prior domain information are independent. Therefore, we integrate them together and make the final estimation of the optimal ED image number n_{ED} by the following cost function,

$$n_{ED} = \arg\max_t \{f(t) \cdot E[p(x_{1:t}|z^h_{1:t}, z^s_{1:t})] \cdot p_0(t)\} \tag{13}$$

where $f(t)$ is the ED distribution density from the ECG signal, $p(x_{1:t}|z^h_{1:t}, z^s_{1:t})$ is the joint posterior density from image analysis, and $p_0(t)$ is the prior density from the domain information.

The clinical protocol requires that the desirable ES should be within the same cardiac cycle of the optimal ED frame. We initially tried to determine the cardiac cycle length by computing the average distance between two consecutive EDs from the ECG signal. Unfortunately, due to the false positives and false negatives of the R-wave detection algorithm, we found that it is difficult to estimate an accurate cardiac cycle length for ES detection. In this paper, we used a different way to solve the problem. The normal heart rate of a human being is between 50 beats per minute (bpm) to 120 bpm. Considering the extremely abnormal cases in the left ventricular analysis, it can be assumed that the heart rate is in a range of 40 to 150 bmp. In other words, one beat lasts 0.4 to 1.5 seconds. With the frame rate f_r, we can easily get the range of a human being's possible cardiac length in terms of frame number $[cMin, cMax]$, where $cMin = 0.4f_r$, $cMax = 1.5f_r$. These two numbers provides us with a reasonable ES search range which covers most of the left ventricular analysis data. In case this is not valid, manual inspection will be suggested to doctors instead. After having the ES search range, we can start from the optimal ED image, and find the first ES image within $[cMin, cMax]$ as the optimal ES image. In other words, the cost function can be defined as follows,

$$n_{ES} = \arg\min_t \{E[p(x_{1:t}|z^h_{1:t}, z^s_{1:t})] \cdot p'_0(t)\} \tag{14}$$

where $p'_0(t)$ is the density of ES search region. It is defined as,

$$p'_0(t) = \begin{cases} \frac{1}{\sqrt{2\pi_{es}}\sigma_{es}} exp\{-\frac{(t-\mu_{es})^2}{2\sigma^2_{es}}\} & \text{if } (n_{ED} + cMin) \leq t \leq (n_{ED} + cMax); \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

where σ^2_{es} is a variance constant, and $\mu_{es} = (cMin + cMax)/2 + n_{ED}$.

6 Experimental Results

We have collected forty six left ventricular angiograms from four hospitals in both United States and Germany. The frame rate was from 12.5 frame per second (fps) to 30 fps.The total frame number to be from 85 to 300 frames, covering 5 to 10 cardiac cycles. Manual inspection was done on all image sequences to label the ground truth of desirable ED and ES images.

Fig. 3 illustrates the results, where the first column shows the data analysis and fusion plots, the second column presents the retrieved ED and ES images, respectively. In Fig. 3(a), the dark blue curve is calculated by the expectation of the histogram-based observation likelihood; the green curve is got from the scene-based likelihood model; the dash blue vertical lines indicate ED candidates estimated from the ECG signal; the black horizontal line is the average of green curve, whose first intersection with the green curve indicates the time when the

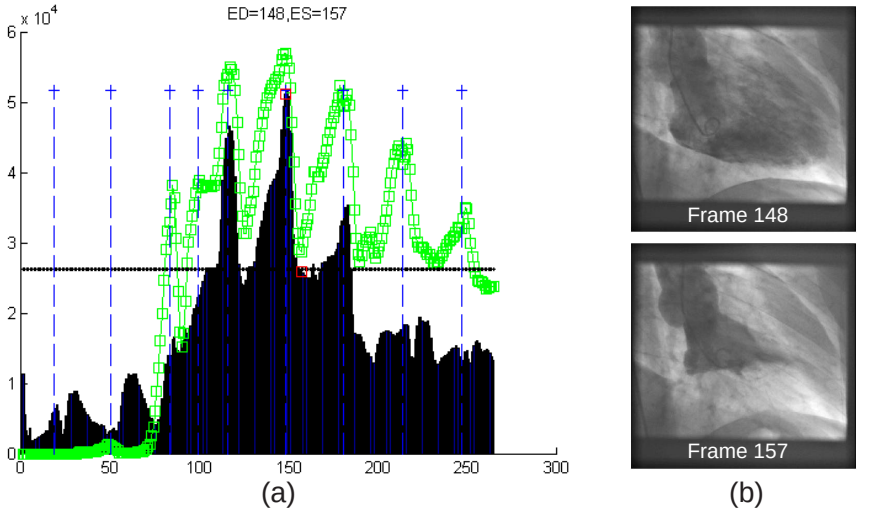


Fig. 3. Retrieval results on real left ventricular angiograms

Table 1. Performance Comparisons between DE-EDES and AutoEDES

Method	DF-EDES	AutoEDES
ED Accuracy	99%	86%
ES Accuracy	95%	72%
Extra-Systolic	Good	None
Low Frame Rate	Good	OK
Background Motion	Good	OK
Lightness Change	Good	Poor
Poor Dye Contrast	Good	Poor
Small ES LV	Good	Poor

dye was injected into the left ventricle; the two red squares show the estimated indices of the optimal ED an ES frames.

In order to demonstrate the robustness of the proposed data fusion based intelligent ED/ES image retrieval approach (DE-EDES), we have compared its performance with our previous reported work AutoEDES in [3]. Table 1 shows the results. As we can see, our new method DF-EDES can achieve very satisfying retrieval results for both ED and ES frames. However, AutoEDES did not work well, which is mainly because we have included more sequences with very poor image quality in the experiments. In order to give a better understanding of advantages and weakness of the two methods, we have also listed a bunch of different tough situations in Table 1. Due to lack of quantitative criteria to

Table 2. Computational Cost Analysis

Speed of Different DF-EDES Modules				AutoEDES
Histogram Model	Scene Model	ECG Processing	Data Fusion	Speed
143 milliseconds	156 milliseconds	23 milliseconds	18 milliseconds	302 milliseconds

describe these situations, we use four qualitative levels "Good", "OK", "Poor", and "None" to show the capability of each approach against the difficulties.

We have implemented the proposed framework in C++ on a laptop with Intel 2.33Ghz Processor and 2G RAM. Without code optimization, the average speed is 250 to 410 milliseconds for the sequences with 85 to 300 frames. Compared with the method in [3], the proposed approach spends additional computational cost on the scene-based image feature extraction, ECG-based R-wave detection, and data fusion, but saves time on adaptive background modeling. Table 2 presents an average comparison of the computational cost between different modules of DF-EDES and AutoEDES method.

7 Conclusions

We have presented a real-time intelligent ED/ES image retrieval framework using Bayesian probability propagation and data fusion. It exploits the advantages of different image features, and integrates both prior domain knowledge and ECG information together to achieve robust ED/ES image retrieval from X-ray left ventricular angiograms.

References

1. Bhargava, V., Hagan, G., Miyamoto, M.I., Ono, S., Ono, S., Rockman, H., Ross Jr., J.: Systolic and Diastolic Global Right and Left Ventricular Functionassessment in Small Animals using An Automated Angiographic Technique. In: Proceedings of Computers in Cardiology, pp. 191–194 (2002)
2. Sandler, H., Dodge, H.T.: The Use of Single Plane Angiocardiograms for the Calculation of Left Ventricular Volume in Man. Amer. Heart J. 75(3), 325–334 (1968)
3. Qu, W., Singh, S., Keller, M.: AutoEDES: A Model-Based BAyesian Framework for Automatic End-Diastolic and End-Systolic Frame Selection in Angiographic Image Sequence. In: Proc. SPIE International Conf. on Medical Imaging (2008)
4. Rabiner, L.R., Juang, B.H.: An Introduction to Hidden Markov Models. IEEE ASSP Mag., 4–15 (1986)
5. Arzeno, N.M., Poon, C.S., Deng, Z.D.: Quantitative Analysis of QRS Detection Algorithms Based on the First Derivative of the Ecg. In: Proc. IEEE Conf. Eng. Med. Biol. Soc., pp. 1788–1791 (2006)

Textile Recognition Using Tchebichef Moments of Co-occurrence Matrices

Marc Cheong and Kar-Seng Loke

School of Information Technology, Monash University Sunway Campus,
Jalan Lagoon Selatan, Bandar Sunway, 46150 Selangor Darul Ehsan, Malaysia
{marc.cheong,loke.kar.seng}@infotech.monash.edu.my

Abstract. The existing use of summary statistics from co-occurrence matrices of images for texture recognition and classification has inadequacies when dealing with non-uniform and colored texture such as traditional 'Batik' and 'Songket' cloth motifs. This study uses the Tchebichef orthogonal polynomial as a way to preserve the shape information of co-occurrence matrices generated using the RGB multispectral method; allowing prominent features and shapes of the matrices to be preserved while discarding extraneous information. The decomposition of the six multispectral co-occurrence matrices yields a set of moment coefficients which can be used to quantify difference between textures. The proposed method have yielded very good recognition rate when used with the *BayesNet* classifier.

Keywords: texture recognition, texture classification, Tchebichef, orthogonal polynomial, co-occurrence matrix, GLCM, textile motifs.

1 Introduction

Batik and *Songket* motifs [1] are traditional Malaysian-Indonesian cloth designs, with intrinsic artistic value and a rich and diverse history. Despite having a history spanning centuries, they are still valued today for their beauty and intricacy, commonplace amongst today's fashion trends.

These patterns and motifs, however, defy a simple means of systematic cataloguing or indexing, and categorization. Linguistic terms are not accurate enough to identify or categorize - with sufficient accuracy - a particular textile motif; save for a few common design patterns, due to the diversity of patterns. Therefore, the pattern identification would have to be by example; making this ideal for content-based image retrieval and recognition.

In this paper, we will be using a collection of traditional *Batik* and *Songket* design motifs as input for performing classification and recognition by extending previous research on texture recognition. The collection consists of 180 different samples [1], sourced from 30 different texture classes (6 samples per class). Refer to Figure 1 for samples of the classes used in this paper.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1017 – 1024, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Fig. 1. Samples of texture motifs from 4 different classes used as sample data for this research

2 Related Work

Grey Level Co-occurrence Matrix or GLCM (also known as Spatial-dependence Matrix) has been known as a powerful method [2] to represent the textures. Textures can be described as patterns of "non-uniform spatial distribution" of grayscale pixel intensities [2]. Allam et. al [3], citing Wezka et al. [4], and Conners and Harlow [5] found that co-occurrence matrices yield better results than other texture discrimination methods. Haralick [6] achieved a success rate of approximately 84% by using the extraction and calculation of summary statistics of the GLCM found in grayscale images, having an advantage in speed compared with other methods [3]. Based on the good acceptance of GLCM approaches to texture recognition, in this research, we have adopted the use of GLCM as the basis for our textile motifs recognition. GLCM-based texture recognition have been used in combination with other techniques, including combining its statistical features with other methods, such as genetic algorithms [7]. Practical applications of GLCM in image classification and retrieval include iris recognition [8], image segmentation [9] and CBIR in videos [10].

For use in color textures, Arvis et al. [11] have introduced a multispectral variation to the GLCM calculation that supports multiple color channels, by separating each pixel's color space into RGB components, and uses pairings of individual color channels to construct multiple co-occurrence matrices. In this paper, we propose using the six RGB multispectral [11] co-occurrence matrices – generated by separating each colored pixel into its Red, Green, and Blue components. RGB color space is selected as opposed to others such as YUV and HSV, as it yields a reasonable [12] rate of success. The orthogonal polynomial moments for these six matrices are used as descriptors for the matrices in place of the summary statistics such as Haralick's measures [2]. Allam et al. [3] have also devised a method using orthonormal descriptors in their work on texture recognition on a 2-class problem, with a less than 2% error rate. Jamil et al. [13, 14] have worked retrieval of *Songket* patterns based on their shapes using geometric shape descriptors from gradient edge detectors. Their method

achieved their best "precision value of 97.7% at 10% recall level and 70.1% at 20% recall level" [13, 14].

3 Description of Current Work

Mathematically, we define the source image as $I(x, y)$, where (x, y) determines the pixel coordinates, and $0 \leq I(x, y) \leq 255$. The multispectral co-occurrence matrix [11] represents the total number of pixel pairs in $I(x, y)$ having a color value i from the color channel a, and the value j from the color channel b. The pixel pairs in may be separated by a vector **T** where:

$$(x_2, y_2) = (t_x+x_1, t_y+y_1) \tag{1}$$

given (x_1, y_1) as coordinate of the first pixel, (x_2, y_2) for the second pixel [3].

We define a co-occurrence matrix of colors a and b ($a, b \in$ {R, G, B}) dealing with pixel pairs in $I(x, y)$ separated by a vector **T** as:

$$C_{abt} = \begin{bmatrix} c_{abt}^{00} & \cdots & c_{abt}^{0j} \\ \vdots & \ddots & \vdots \\ c_{abt}^{i0} & \cdots & c_{abt}^{ij} \end{bmatrix} \tag{2}$$

We then define $c_{abt}(i_a, j_b)$ as:

$$c_{abt}(i_a, j_b) = \sum_{x,y} \sum_{tx,ty \in U} \delta[I(x, y) - i] \times \delta[I(x + t_x, y + t_y) - j] \tag{3}$$

given: i_a and j_b are intensity values from channels a and b respectively, **T** is the distance vector between the two pixels), δ the Kronecker Delta, and $x, y \in I$.

The set U of all possible t_x and t_y values satisfy the condition $x^2 + y^2 = r^2$; r being a fixed distance from the center pixel, $r \in \mathbf{Z}$; yielding a co-occurrence matrix with rotation-invariance.

Each of the six individual multispectral matrices, C_{ab} ($a, b \in$ {R, G, B}) is converted to a grayscale image, G_{ab}(i, j), such that $0 \leq i, j \leq 255$ (see Figure 1). The pixel intensity at any given position (i, j) correlates directly with the value in the co-occurrence matrix $C_{ab}(i, j)$, through the following equation:

$$g_{ab}(i, j) = \frac{c_{ab}(i, j)}{\max(c_{ab}(i, j))} \times 255 \tag{4}$$

After conversion $\min(c_{ab}(i, j))$ will have $g_{ab} = 0$, while $\max(c_{ab}(i, j))$ has $g_{ab} = 255$. However, outlying values (i.e. small values in C_{ab}) that contribute to the overall shape of the co-occurrence matrix will be lost during decomposition into moments. To solve this problem, the output image is then histogram-equalized to highlight such outlying values. By visual inspection of the generated matrices, images in the same texture class will have a similar set of the six matrices. For example, notice the similarity between the first two representations of multispectral matrices (samples of the same texture class) as opposed to other matrices (samples from different classes) in Figure 2.

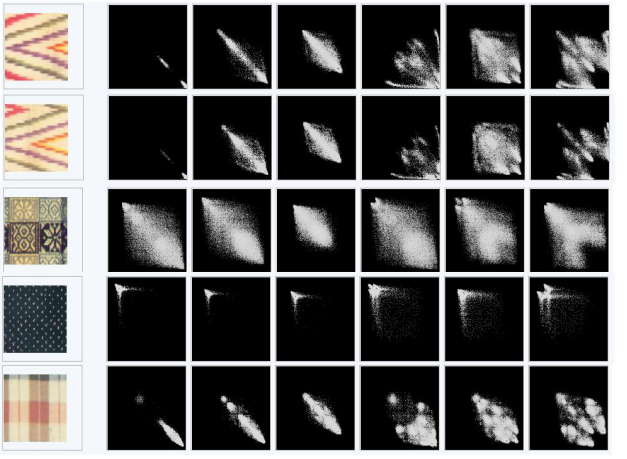


Fig. 2. Multispectral RGB co-occurrence matrices for 'Batik' motifs. Each row shows: 'Batik' motif and its corresponding matrices from the RR, GG, BB, RG, GB, and BR channels.

Current methods (using second-order statistics of co-occurrence matrices) compress the features of the matrix into a set of 'summary statistics' summarizing important textural features – Haralick [6] identified thirteen of them; five are commonly used [10, 11]. We propose to use the "shape" information from the multispectral matrices as a basis in texture recognition and classification. Such "shape" information can be seen as capturing more complete information rather than information such its skew, contrast, etc.

We introduce the usage of orthogonal polynomials as a means of representing the information found in the co-occurrence matrices. See et al. [15] have shown that discrete orthogonal polynomials such as the Tchebichef discrete orthogonal polynomial can be an effective way of representing any 2D function. Various orthogonal polynomial moments, such as Zernike [16] and Hermite [17] have been applied to texture classification. However, our approach differs in that we apply the orthogonal polynomial moments on the co-occurrence matrix image, not on the image directly. Our approach require that the multispectral co-occurrence matrices to be treated as an image, and hence can be represented as a series of image moments [12]. The limited finite expansion of the moments allow only prominent features to be preserved while discarding those moments which carry little or no information [15, 18]. The first few moments encode gross overall shape and other moments carry finer details; thus, by discarding higher moments, we are able to save on complexity while preserving the entire set of second-order textural statistics in the multispectral matrix.

This paper applies the Tchebichef method to decompose the six generated multispectral co-occurrence matrices and using the resulting moment values as basis for texture discrimination. Using the findings from See et al. [15], we can decompose the generated visual representation of the 6 multispectral matrices into a number of coefficients.

Mathematically, the following decomposition function transforms the matrices' image intensity into moment orders M_{pq} [15]:

$$M_{pq} = \frac{1}{\rho(p)\rho(q)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} m_p(x) w(x) m_q(y) w(y) f(x,y) \tag{5}$$

where: $0 \leq p, q, x, y \leq N$-1; $m_n(x)$ is a set of finite discrete orthogonal polynomials, $w(x)$ the weight function, and $\rho(n)$ the rho function.

The Tchebichef polynomial is defined mathematically as [14]:

$$m_n(x) = n! \sum_{k=0}^{n} (-1)^{n-k} \binom{N-1-k}{n-k} \binom{n+k}{n} \binom{x}{k} \tag{6}$$

$$\rho(n) = (2n)! \binom{N+n}{2n+1} \tag{7}$$

$$w(x) = 1 \tag{8}$$

given: m_n is the n-th Tchebichef polynomial, ρ(n) the rho function and w(x) the weight function. Let N be the number of moment orders used for the decomposition process. The total number of coefficients resulting from the decomposition process for each matrix is N^2**.** For the 6 matrices involved, the total number of coefficients per sample image is therefore $6(N)^2$. These are generated from the database of 180 sample images from 30 classes of 'Batik' and 'Songket' textures. The coefficients are then stored in CSV format and imported into Weka [19] for further analysis.

In Weka, two unsupervised clustering algorithms and two supervised classifiers are used to classify our sets of generated moment coefficients. The unsupervised clusterers are *IBk* (*k*-means with the *k* value set to the number of expected classes, i.e. 30), *FarthestFirst* (an optimized implementation of the *k*-means method); while the two supervised classifiers are *BayesNet* and *kNN* (*k*-nearest neighbor, with the *k* value set to 5). All of them use default parameters as defined in Weka. For the supervised classifier, we use 10-fold cross-validation to automatically partition the test and training data: the collection of sample data is partitioned into 10 mutually-exclusive partitions (called folds) [20].

The k-means algorithm by McQueen [21] works to partition our sample data (unsupervised) into *k* distinct clusters. The naïve K-means algorithm does so by minimizing total intra-cluster variance; in the context of our methods, it tries to identify the samples which minimize the variance within a particular texture class, thereby properly grouping these samples by texture class. *FarthestFirst* [19] is an implementation of an algorithm by Hochbaum and Shmoys [19], cited in Dasgupta and Long [22]. It works "as a fast simple approximate clusterer" modeled after the naïve k-means algorithm. kNN (the k-nearest neighbor) classifier works by assigning a texture (whose class is yet unknown) to the class in which the majority of its *k* neighbors belong to. In this case, we compare the linear distance between a texture sample and each of its *k* (we fix the value of k=5) neighbors, finally assigning it a class based on the majority of its 5 neighbors. The *BayesNet* Bayesian network learning algorithm in Weka uses the K2 hill-climbing strategy to construct a Bayesian network from the given coefficient data; by constructing a model to determine Bayesian probability of a single sample image as belonging to a class.

4 Experimental Results

For the purposes of this paper, the Tchebichef polynomial decomposition is performed with moment orders, N = 10, as it was previously determined to have the best degree of accuracy with the least processing time [23]. The following table shows the experimental results obtained in Weka [18].

Table 1. Experimental results as determined in Weka for each of the four methods

Method	Samples	Correct	Incorrect	Percentage
Supervised: *BayesNet*	180	179	1	99.44%
Supervised: 5NN (*kNN*)	180	176	4	97.78%
Unsupervised: *FarthestFirst*	180	173	7	96.11%
Unsupervised: k-means (*IBk)*	180	167	13	92.78%

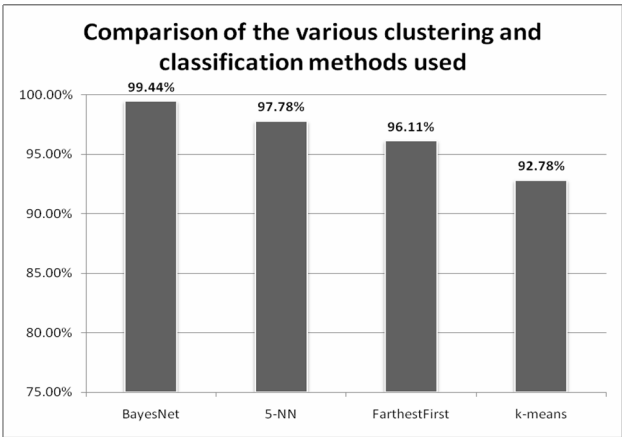


Fig. 3. Graph comparing the correct classification percentage for each of the four methods used

5 Discussion and Analysis

Prior research on the GLCM has focused predominantly on textures. Arvis et al. [11] with their multispectral co-occurrence matrix method, with a 5-Nearest Neighbors classifier yielding a 97.9% percentage of good classification for VisTex [24] textures. Previous research work involving color texture analysis using a combination of Gabor filtering and the multispectral method on the Outex [25] database has yielded a rate of success of 94.7% [25]. Allam's result of a 2% error rate [3] differs in the fact it is only applied to a 2-class problem, restricted to grayscale texture. This differs in our motivation of using the "shape" of the co-occurrence pattern.

The results for 'Batik' and 'Songket' achieved here are among the best for such kinds of textile patterns based on the limited prior research found [13, 14]. Experimental tests on co-occurrence matrices using summary statistics suggest that summary statistics may

not always capture the full representation of the co-occurrence matrix: the rationale being many similar distributions having the possibility of producing a similar value [26] and resulting in a lower success rate [6]. Design motifs such as those found in textiles tend to have a more non-uniform distribution in the GLCM as opposed to textures. This also makes it difficult to be captured by Haralick's summary statistics [6] as the "shape" information is not adequately represented. Our method has the best success rate using the Tchebichef orthogonal polynomial, with 10 order of moments used [23]. This is due to the fact that with Tchebichef, the reconstructed matrices strike a balance between preserving the shape of the matrices' visual representation and a good degree of variance when matching with other samples.

6 Conclusion

We have successfully demonstrated the multispectral co-occurrence matrices method for use in the recognition of *Batik* and *Songket* design motifs and introduced the use of the Tchebichef orthogonal polynomial to decompose each of these matrices into a series of moments as a means to capture more complete second-order pixel statistics information. The advantage to this method is having a good degree of accuracy as compared to the use of summary statistics which is commonly used in GLCM research. We have also shown that this method is viable in matching non-uniform design motifs as opposed to only textures. This makes our approach suitable to be used in image retrieval applications for not only traditional *Batik* and *Songket* textile motifs but other design motifs.

References

1. Ismail, S.Z.: Malay Woven Textiles: The Beauty of a Classic Art Form. Dewan Bahasa dan Pustaka (1997)
2. Davis, L.S.: Image Texture Analysis Techniques - A Survey. In: Simon, J.C., Haralick, R.M. (eds.) Digital Image Processing, D. Reidel, Dordrecht (1981)
3. Allam, S., Adel, M., Refregier, P.: Fast Algorithm for Texture Discrimination by Use of a Separable Orthonormal Decomposition of the Co-occurrence Matrix. Applied Optics 36, 8313–8321 (1997)
4. Weszka, J.S., Dyer, C.R., Rosenfeld, A.: A Comparative Study of Texture for Terrain Classification. IEEE Trans. on Sys., Man, and Cyber. 6, 265–269 (1976)
5. Conners, R.W., Harlow, C.A.: A Theoretical Comparison of Texture Algorithms. IEEE Trans. on PAMI 3, 204–222 (1980)
6. Haralick, M., Shanmugam, K., Dinstein, I.: Textural Features for Image Classification. IEEE Trans. on Sys., Man, and Cyber. 3, 610–621 (1973)
7. Walker, R.F., Jackway, P.T., Longstaff, I.D.: Recent Developments in the Use of the Co-occurrence Matrix for Texture Recognition. In: DSP 1997: 13th International Conf. on DSP, vol. 1, pp. 63–65 (1997)
8. Zaim, A., Sawalha, A., Quweider, M., Iglesias, J., Tang, R.: A New Method for Iris Recognition Using Gray-level Coccurence Matrix. In: IEEE International Conf. on Electro/Information Technology, pp. 350–353 (2006)
9. Abutaleb, A.S.: Automatic Thresholding of Gray-level Pictures Using Two-dimensional Entropies. Computer Vision Graphics Image Processing 47, 22–32 (1989)

10. Kim, K., Jeong, S., Chun, B.T., Lee, J.Y., Bae, Y.: Efficient Video Images Retrieval by Using Local Co-occurrence Matrix Texture Features and Normalised Correlation. In: Proceedings of The IEEE Region 10 Conf., vol. 2, pp. 934–937 (1999)
11. Arvis, V., Debain, C., Berducat, M., Benassi, A.: Generalization of the Cooccurrence Matrix for Colour Images: Application to Colour Texture Classification. Image Analysis and Stereology 23, 63–72 (2004)
12. Chindaro, S., Sirlantzis, K., Deravi, F.: Texture Classification System using Colour Space Fusion. Electronics Letters 41 (2005)
13. Jamil, N., Bakar, Z.A., Sembok, T.M.T.: Image Retrieval of Songket Motifs using Simple Shape Descriptors. In: Geometric Modeling and Imaging– New Trends (GMAI 2006) (2006)
14. Jamil, N., Bakar, Z.A.: Shape-Based Image Retrieval of Songket Motifs. In: 19th Annual Conference of the NACCQ, pp. 213–219 (2006)
15. See, K.W., Loke, K.S., Lee, P.A., Loe, K.F.: Image Reconstruction Using Various Discrete Orthogonal Polynomials in Comparison with DCT. Applied Mathematics and Computation 193, 346–359 (2008)
16. Wang, L., Healey, G.: Using Zernike Moments for the Illumination and Geometry Invariant Classification of Multispectral Texture. IEEE Trans. on Image Processing 7, 196–203 (1998)
17. Krylov, A.S., Kutovoi, A., Leow, W.K.: Texture Paramaterization with Hermite Functions. Computer Graphics and Geometry 5, 79–91 (2003)
18. Kotoulas, L., Andreadis, I.: Image Analysis Using Moments. In: 5th Int. Conf. on Tech. and Automation, pp. 360–364 (2005)
19. The University of Waikato: Weka 3, http://www.cs.waikato.ac.nz/ml/weka/
20. Kohavi, R.: A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection. In: Fourteenth International Joint Conference on Artificial Intelligence, San Mateo, CA, pp. 1137–1143 (1995)
21. MacQueen, J.B.: Some Methods for Classification and Analysis of Multivariate Observations. In: 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297. University of California Press, Berkeley (1967)
22. Dasgupta, S., Long, P.M.: Performance Guarantees for Hierarchical Clustering. Journal of Computer and System Sciences 70, 555–569 (2005)
23. Cheong, M., Loke, K.S.: An Approach to Texture-Based Image Recognition by Deconstructing Multispectral Co-occurrence Matrices using Tchebichef Orthogonal Polynomials (2008)
24. MIT Media Lab: Vision Texture Database, http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html
25. Ojala, T., Maenpaa, T., Pietikainen, M., Viertola, J., Kyllonen, J., Huovinen, S.: Outex - new Framework for Empirical Evaluation of Texture Analysis Algorithms. In: 16th International Conference on Pattern Recognition, vol. 1, pp. 701–706 (2002)
26. Walker, R.F., Jackway, P.T., Longstaff, I.D.: Improving Co-occurrence Matrix Feature Discrimination. In: DICTA 1995, pp. 643–648 (1995)

Generic Summarization Using Non-negative Semantic Variable

Sun Park

Department of Computer Engineering, Honam University, Korea
sunpark@honam.ac.kr

Abstract. The methods using supervised algorithms for generic document summarization are time-consuming because they need a set training data and associated summaries. We propose a new unsupervised method using the Non-negative Semantic Variable to select the sentences for automatic generic document summarization. The proposed method selects meaningful sentences for generic document summarization. Besides, it can improve the quality of generic summaries because the extracted sentences are well covered with the major topics of document. And also it doesn't need a set training data because it is an unsupervised method. The experimental results demonstrate that the proposed method achieves better performance the other method.

1 Introduction

Generic document summarization is the process of distilling the most important information from a source to produce an abridged version for a particular user and task [11]. Document summaries can be either generic summaries or query-focused summaries. A generic summary presents an overall sense of the documents' contents. A query-focused summary presents the contents of the document that are related to the user's query [10, 13].

There are two methods for automatic generic document summarization: supervised or unsupervised methods. The supervised methods typically make use of human-made summaries or extracts to find features or parameters of summarization algorithms, while unsupervised approaches determine relevant parameters without regard to human-made summaries [12].

The recent studies for generic document summarization based on supervised algorithms are as follows: Chuang and Yang proposed a method that uses a supervised learning algorithm to train the summarizer to extract important sentence segments that are represented by a set of predefined features [3]. The methods to use supervised learning algorithms requires a set of training documents and associated summaries. It is time consuming and irrelevant for many applications because they require labeling large amount of text spans for training summarization systems. Amini and Gallinari proposed a method that uses a semi-supervised algorithm for training classification models for text summarization. These algorithms make use of few labeled data together with a larger amount of unlabeled data [1].

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1025 – 1031, 2008.
© Springer-Verlag Berlin Heidelberg 2008

The unsupervised methods for generic document summarization are as follows: Gong and Liu proposed the method using the Latent Semantic Analysis (LSA) technique to semantically identify important sentences for summary creation [5]. Sum and Shen proposed the method that uses Adapted LSA to leverage the clickthrough data for Web-page summarization [13]. The LSA was used as the basis for sentence selection of a given document, using the components of multiple singular vectors. The singular vectors other than the one corresponding to the largest singular value can have both positive and negative components, making ranking sentences by singular vector component value less meaningful [15]. Nomoto and Matsumoto proposed the method that uses a variation of the K-means method to cluster the sentences of a document into different topical groups, and then apply a sentence weighting model within each topical group for sentence selection [12]. Zha proposed the method that clusters sentences of a document into topical groups and then, within each topical group, to select the key-phrases and sentences by their saliency scores [15]. Harabagiu and Lacatusu described representations of five different topics and introduced a novel representation of topics based on topic themes [6]. The computational complexities of topic based methods are very high because they require many preprocessing steps to find topics from given documents.

Unlike the LSA, the NMF can find a parts representaion of the data because non-negative constraints are compatible with the intuitive notions of combining parts to form a whole, which is how NMF learns a parts-based representation. NMF is to decompose a non-negative matirx into Non-negative Feature Matrix(NFM) and Non-negative Variable Matrix(NVM) [8,9,14].

In this paper, we propose the new method that makes generic summaries by extracting sentences using non-negative semantic variable. The proposed method in this paper has the following advantages: First, it selects meaningful sentences for text summarization, because the NMF uses non-negative semantic feature vectors to find meaningful sentences. Second, it can improve the quality of generic summaries because the extracted sentences are well covered with the major topics of document. Third, it is to unsupervised text summarization that not requires a training set of documents and relevant extract summaries. Also, it has the computation of low cost and can select sentences easily.

The rest of the paper is organized as follows: Section 2 describes the proposed text summarization method and in section 3 describes the performance evaluation. We conclude the paper in section 4 with future research.

2 Generating Generic Summaries

There are several topics in a document. Many sentences support some major topics to form the major content of the document. Other sentences are described to supplement less important topics. So, the major topics must be covered as much as possible to get a better generic summary, and at the same time, keep redundancy to a minimum [5].

In this section, we propose a method that creates generic document summaries by selecting sentences using the semantic variable from NMF. The proposed method consists of the preprocessing step and the summarization step. In the following subsections, the two steps are described in detail.

2.1 Preprocessing

In the preprocessing step, after a given English document is decomposed into individual sentences, we remove all stopwords by using Rijsbergen's stopwords list and perform words stemming by Porter's stemming algorithm [4]. Unlike English documents, Korean documents are preprocessed using HAM (Hangul Analysis Module) which is a Korean language analysis tool based on Morpheme analyzer [7]. Then we construct the weighted term-frequency vector for each sentence in the document by Equation (1) [2, 3, 4, 7].

Let $T_i = [\ t_{1i}, t_{2i}, \ldots, t_{ni}\ ]^T$ be the term-frequency vector of sentence i, where elements t_{ji} denotes the frequency in which term j occurs in sentence i. Let $\boldsymbol{A}$ be $m \times n$ weighted terms by sentences matrix, where m is the number of terms and n is the number of sentences in a document. Let element A_{ji} be the weighted term-frequency of term j in sentence i.

$$\mathrm{Aji} = \mathrm{L(j,\ i)} \cdot \mathrm{G(j,\ i)} \tag{1}$$

Where $L(j, i)$ is the local weighting for term j in sentence i, and $G(j, i)$ is the global weighting for term j in the whole documents. That is,

1. No weight: $L(j, i) = t_{ji}$, $G(j, i) = 1$ for any term j in sentence i.
2. Augmented weight: $L(j, i) = 0.5 + 0.5 * (t_{ji} / tfmax)$, $G(j, i) = \log(N/n(j))$.

Where $tfmax$ is the frequency of the most frequently occurring term in the sentecnce, N is the total number of sentences in the document, and $n(j)$ is the number of sentences that contain term j [5].

2.2 Generic Document Summarization Using Non-negative Semantic Variable

The summarization step selects sentences for generic summarization by using the Non-negative Semantic Variable (NSV) from NMF. We perform the NMF on $\boldsymbol{A}$ to obtain the two non-negative semantic feature matrix $\boldsymbol{W}$ and non-negative semantic variable matrix $\boldsymbol{H}$ such that:

$$A \approx WH \tag{2}$$

$\boldsymbol{W}$ is an $m \times r$ matrix and $\boldsymbol{H}$ is an $r \times n$ matrix. Usually r is chosen to be smaller than n or m, so that $\boldsymbol{W}$ and $\boldsymbol{H}$ are much smaller than the original matrix $\boldsymbol{A}$. This results in a compressed version of the original data matrix. We keep updating $\boldsymbol{W}$ and $\boldsymbol{H}$ until $\left\|A - WH\right\|^2$ converges under the predefined threshold. The update rules are as follows [8, 9, 14]:

$$H_{\alpha\mu} \leftarrow H_{\alpha\mu} \frac{(W^T V)_{\alpha\mu}}{(W^T WH)_{\alpha\mu}}, \qquad W_{i\alpha} \leftarrow W_{i\alpha} \frac{(VH^T)_{i\alpha}}{(WHH^T)_{i\alpha}} \tag{3}$$

A column vector corresponding to j'th sentence $A_{\bullet j}$ can be represented as a linear combination of semantic feature vectors $W_{\bullet l}$ and semantic variable h_{lj} as follows:

$$A_j = \sum_{l=1}^{r} h_{lj} W_{\bullet l} \tag{4}$$

Where $W_{\bullet l}$ is the l'th column vector of $\boldsymbol{W}$.

The powers of the two non-negative matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ are described as follows: All semantic variables (h_{lj}) are used to represent each sentence. $\boldsymbol{W}$ and $\boldsymbol{H}$ are represented sparsely. Intuitively, it make more sense for each sentence to be associated with some small subset of a large array of topics ($W_{\bullet l}$), rather than just one topic or all the topics. In each semantic feature ($W_{\bullet l}$), the NMF has grouped together semantically related terms. In addition to grouping semantically related terms together into semantic features, the NMF uses context to differentiate between multiple meanings of the same term [8].

We propose a method to select sentences based on the semantic variable by NMF. We define the *semantic weight of a sentence weight( )* as follows:

$$weight(H_{i\bullet}) = \sum_{q=1}^{n} H_{iq} \tag{5}$$

The weight ($H_{i\bullet}$) means the relative relevance of i'th semantic feature ($W_{\bullet i}$) among all semantic features. The generic relevance of a sentence means how much the sentence reflects major topics which are represented as semantic features.

The proposed algorithm for generic document summarization is as follows:

1. Decompose the document $\boldsymbol{D}$ into individual sentences, and let k be the numb er of sentences for generic document summarization.
2. Perform the stopwords removal and words stemming operations.
3. Construct the weighted terms by sentences matrix $\boldsymbol{A}$ using Equation (1).
4. Perform the NMF on the matrix $\boldsymbol{A}$ to obtain the matrix $\boldsymbol{H}$ using Equation (3)
5. for each sentence j
   calculate the *semantic weight of sentence j* $weight(H_{j\bullet})$
6. select k sentences with highest semantic weight values

3 Performance Evaluation

We conducted the performance evaluation using the Reuters-21578 [16] document corpora and Yahoo-Korea news [17]. Three independent evaluators were employed to manually create summarization on the 200 document from Reuters-21578 corpora and Yahoo-Korea news. Each document in Reuters-21578 has an average of 2.67 sentences selected by evaluators. Also, each document in Yahoo Korea news has an average of 2.08 sentences selected by evaluators. Table 1 provides the particulars of the evaluation data corpus.

Table 1. Particulars of the Evaluation Data Corpus

Document attributes	Reuters-21578	Yahoo-Korea News
Number of docs	100	100
Number of docs with more than 10 sentences	35	32
Avg sentences / doc	10.09	10.1
Min sentences / doc	3	3
Max sentences / doc	40	36

We used the recall (*R*), precision (*P*), along with *F* measure to compare the performances of the two summarization methods, LSA and NSV. The LSA denotes Gong's method [5]. The NSV denotes the proposed method. The NoWeight uses the No weight. The Weight uses the Augmented weight. Let S_{man}, S_{sum} be the set of sentences selected by the human evaluators, and the summarizer, respectively. The standard definitions of R, P, and F are defined as follows [4, 7]:

$$R = \frac{|S_{man} \cap S_{sum}|}{|S_{man}|} \qquad P = \frac{|S_{man} \cap S_{sum}|}{|S_{sum}|} \qquad F = \frac{2RP}{R+P} \tag{5}$$

In this paper, we conducted the performance evaluation on the two methods using the same data and Equation (5). The average number of sentences selected by human evaluators is 2.55 on both the LSA and the proposed method. The evaluation results are shown in Figure 1 and Figure2.

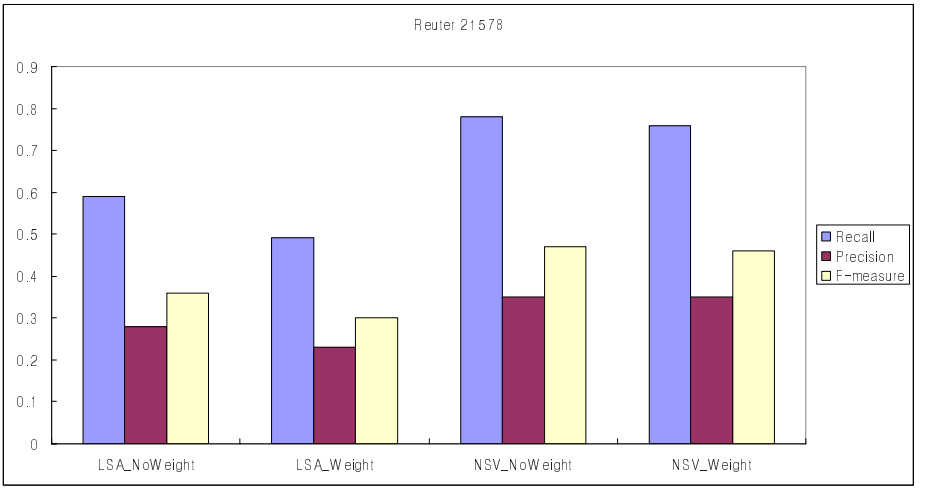


Fig. 1. Evaluation Results using Reuter 21578

Experimental results show that the proposed method surpasses LSA. Our method uses NSV to find the semantic feature with non-negative values for meaningful summarization [8, 9, 14].

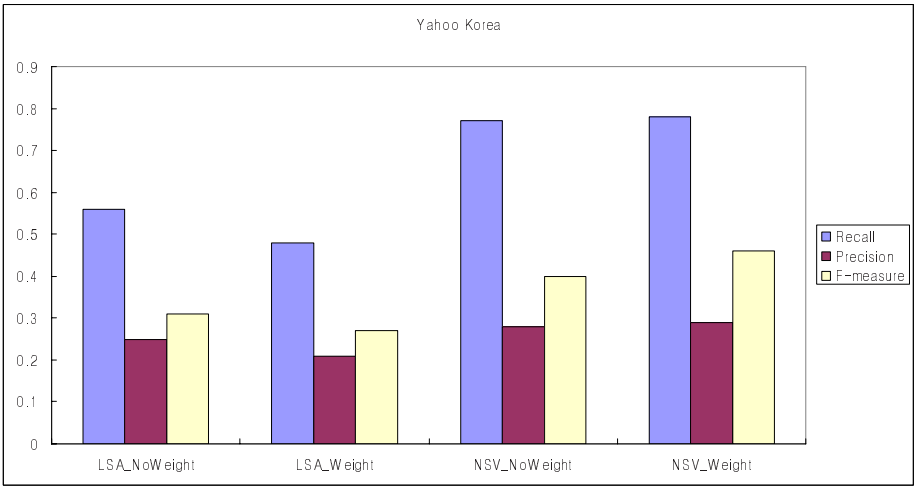


Fig. 2. Evaluation Results using Yahoo Korea

4 Conclusions and Future Work

In this paper, we have presented a generic document summarization method using the *semantic weight of a sentence* based on the non-negative semantic variable from NMF. The proposed method in this paper has the following advantages. it selects more meaningful sentences for generic document summarization than that using the LSA because the NMF has the better power to grasp the innate structure of a document. Besides, it can improve the quality of generic summaries because the extracted sentences are well covered with the major topics of document. And its selection steps are simpler. Moreover, it is an unsupervised method that does not require a set of training data.

In future work, we have a plan to enhance our method by considering other weighting schemes and the correlation between the semantic features and semantic variables. Moreover, we will study the method to select automatically the number of selected sentences by considering the size of a document.

References

1. Amini, M.R., Gallinari, P.: The Use of Unlabeled Data to Improve Supervised Learning for Text Summarization. In: Proceeding of ACM SIGIR 2002, pp. 105–112 (2002)
2. Chakrabarti, S.: Mining the Web: Discovering Knowledge from Hypertext Data. Morgan Kaufmann, San Francisco (2003)
3. Chuang, W.T., Yang, J.: Extracting Sentence Segments for Text Summarization: A Machine Learning Approach. In: Proceeding of ACM SIGIR 2000, pp. 152–159 (2000)
4. Frankes, W.B., Baeza-Yaes, R.: Information Retrieval: Data Structure & Algorithms. Prentice-Hall, Englewood Cliffs (1992)
5. Gong, Y., Liu, X.: Generic Text Summarization Using Relevance Measure and Latent Semantic Analysis. In: Proceeding of ACM SIGIR 2001, pp. 19–25 (2001)

6. Harabagiu, S., Finley, L.: Topic Themes for Multi-Document Summarization. In: Proceeding of ACM SIGIR 2005, pp. 202–209 (2005)
7. Kang, S.S.: Information Retrieval and Morpheme Analysis. HongReung Science Publishing Co. (2002)
8. Lee, D.D., Seung, H.S.: Learning the Parts of Objects by Non-negative Matrix Factorization. Nature 401, 788–791 (1999)
9. Lee, D.D., Seung, H.S.: Algorithms for Non-negative Matrix Factorization. Advances in Neural Information Processing Systems 13, 556–562 (2001)
10. Marcu, D.: The Automatic Construction of Large-scale Corpora for Summarization Research. In: Proceeding of ACM SIGIR 1999, pp. 137–144 (1999)
11. Mani, I., Maybury, M.T.: Advances in Automatic Text. The MIT Press, Cambridge (1999)
12. Nomoto, T., Matsumoto, Y.J.: A New Approach to Unsupervised Text Summarization. In: Proceeding of ACM SIGIR 2001, pp. 26–34 (2001)
13. Sum, J.T., Shen, D., Zeng, H.J., Yang, Q., Lu, Y., Chen, Z.: Web-Page Summarization Using Clickthrough Data. In: Proceeding of ACM SIGIR 2005, pp. 194–201 (2005)
14. Xu, W., Liu, X., Gong, Y.: Document Clustering Based on Non-negative Matrix Factorization. In: ACM SIGIR, Toronto, Canada (2003)
15. Zha, H.Y.: Generic Summarization and Keyphrase Extraction Using Mutual Reinforcement Principle and Sentence Clustering. In: Proceeding of ACM SIGIR 2002, pp. 113–120 (2002)
16. `http://kdd.ics.uci.edu/databases/reuters21578/reuters21578.html` (2008)
17. `http://kr.news.yahoo.com/` (2008)

An Image Indexing and Searching System Based Both on Keyword and Content

Nan Zhang[1] and Yonghua Song[2]

[1] Department of Computer Science and Software Engineering,
Xi'an Jiaotong-Liverpool University
[2] Department of Electrical and Electronic Engineering,
Xi'an Jiaotong-Liverpool University
{nan.zhang,yonghua.song}@xjtlu.edu.cn

Abstract. Content-based image retrieval (CBIR) has certain advantages over those pure keyword-based. CBIR indexes images by visual features that are extracted from the images. This may save the effort spent on the manual annotation. However, because low-level visual features, such as colour and texture, often carry no high-level concepts, images retrieved purely based on content may not match with the intention of the user. The work presented in this paper is an image retrieval system that bases both on text annotations and visual contents. It indexes and retrieves images by both keywords and visual features, with the purpose that the keywords may mend the gap between the semantic meaning an image carries and its visual content. Tests were made on the system that have demonstrated that such a hybrid approach did improve retrieval precisions over those pure content-based.

Keywords: Content-based image retrieval, Keyword-based image retrieval, Hybrid image retrieval, MPEG-7.

1 Introduction

Content-based image retrieval (CBIR) was first proposed in the early 1990's with the purpose of addressing the two main problems intrinsically associated with the traditional keyword based image approaches [1]. In pure keyword based image retrieval systems, images are annotated by keywords that reflect the meanings carried by the image, and then are retrieved according to the matching of the keywords in the query with those found with the images. However, together with the growth of the number of images in a collection up to huge numbers, the amount of labor required by the manual labelling exploded beyond measure.

CBIR was first proposed with the purpose of relieving the labor required by the manual annotations. Unlike its keyword based counterpart, the approach of CBIR extracts information from the image itself and then uses that information as the basis of indexing and comparison. Such often used information include various calculations on image colour, transformations on texture, different representation schemes on shape, and spatial related descriptors. In about 2000,

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1032–1039, 2008.
© Springer-Verlag Berlin Heidelberg 2008

ISO released the MPEG-7 standard which described representation methods for colour, texture and shape [2]. However, over the years, researchers gradually realised that there is a gap between those low-level feature descriptors and the high-level semantic meaning that an image carries. Those colour, texture, and shape descriptors by themselves do not have indications on the meaning of an image. They do not tell us what an image describes.

Because of that gap, in recent years, much research efforts have been shifted from designing sophisticated content extracting algorithms to bringing close the relation between the low-level content descriptors and high-level semantic meanings. We may list two technical streams among such efforts. One is to enrich the formations of query, such as allowing the user to query an image retrieval system by sketch [3]. The other stream is to provide mechanisms of relevance feedback to refine the retrieved results and therefore to enhance accuracy and user experience [4].

The system that we present in this paper is a hybrid image retrieval system in the sense that the indexing and searching are based on both keywords and visual features such as colour and texture. We would argue that such an approach may benefit from both the content-based and keyword-based technologies.

Organisation of the paper. Related work is discussed in Section 2. An overview of the system is presented in Section 3. The keyword-based retrieval in our system is discussed in Section 4. The content-based retrieval in our system is discussed in Section 5. Tests and their results are presented in Section 6. Conclusions are drawn in Section 7, as well as the future studies are outlined.

2 Related Work

In CBIR, colour and texture features used to represent images were extensively studied [5,6]. The colour descriptors selected by the MPEG-7 standard [2] are *dominant colour*, *scalable colour*, *colour structure* and *colour layout*. The *colour auto correlogram* descriptor is a simplified version of colour correlogram due to the latter's computational expensiveness. MPEG-7 proposed three texture descriptors: *texture browsing*, *homogeneous texture* and *local edge histogram*. The colour and texture descriptors employed by our system are colour layout, colour auto correlogram and local edge histogram.

Caliph & Emir are open-source tools written in Java for image annotation and retrieval [7]. Caliph extracts visual features from an image using the MPEG-7 descriptors – colour layout, scalable colour, edge histogram and dominant colour. Emir is the retrieval counterpart of Caliph. Emir creates Lucene index [8] on the text annotations and visual descriptors saved in the XML files and supports retrieving both by image examples and by keywords. The Lire (Lucene Image REtrieval) library was created as a part of the Caliph & Emir project, from which the code of the CBIR part of our system was derived.

Image retrieval systems that unify keywords and visual contents can be found in the literature, such as [9]. A relevance feedback scheme was employed in their system.

3 System Overview

The system we have created is an image indexing and retrieval application entirely written in Java. Conceptually, the system can be viewed as consisting of three parts:

- the part dealing with keyword-based full-text indexing and retrieval,
- the part dealing with content-based indexing and retrieval,
- and, the part dealing with interactions with the user.

The system must have the support from a database management system, for which we have used Microsoft SQL Server 2000 (with service pack 4). Index catalogs are created on the descriptions of images by the full-text index facility from the SQL Server. Keyword-based comparisons are made through the execution of SQL statements.

The CBIR part contains an index creator and an image searcher. The index creator is a Java Swing based tool that extracts from a JPEG image the three colour and texture features employed in our system, and saves them in a Lucene index, with each group of the three visual features representing an image from the collection. The image searcher provides distance measuring methods for each of the visual descriptors. The visual features extracted from the user-uploaded example image are compared linearly to every group of the three visual features saved in the Lucene index. A relevance factor is calculated that measures the distance between the example image and each of the images represented by the corresponding group of the visual descriptors. The images in the collection are then sorted in descending order by the values of the relevance factors, and the most relevant are returned as the result.

The web-based part that interacts with the user provides a JSP page that allows the user to upload a sample image and to specify some keywords. Upon receiving a request the servlet will start searching through the full-text catalog, trying to match up the terms saved in the index with the keywords. The content-based searching will be activated as well to look into the Lucene index. The images returned to the user is the intersection between the two sets found by the two methods. The hope is that the returned images not only bear similar visual effects with the sample image, but also carry close concept.

4 Keyword-Based Retrieval

We shall have more detailed explanations on the keyword-based retrieval in this section and on the content-based retrieval in the next.

Information about all the images from the collection are stored in a local database hosted by the SQL Server 2000. The table that contains the information has three fields – “Type”, “Filename” and “Description”. The type field denotes whether this record is about an image file, a video file, or an audio file. (Besides the images, we have videos and audios in the collection. But we retrieve only the images.) The file-name field records the name of the media file without leading

directories. This is to allow the media files to be moved to different folders. The description field contains text phrases that characterise the media file. Full-text index catalog is created in this field. As a requirement of the SQL Server to establish the full-text index, the file-name field is chosen as the primary key of the table. The searchings are made in the created full-text index catalog through the executions of SQL statements.

5 Content-Based Retrieval

The index creator extracts from the images the visual features represented by the colour layout descriptor and the edge histogram descriptor taken from MPEG-7, and another colour auto correlogram descriptor. The descriptors are then saved into a Lucene index established in the local file system.

5.1 The Colour Layout Descriptor

The colour layout descriptor [10] is a colour descriptor defined on the YCrCb space. To extract a colour layout descriptor, the image is first partitioned into 8×8 blocks. Following that, dominant colours are computed to represent each of the blocks. An 8×8 discrete cosine transform (DCT) is then performed on the Y, Cb and Cr channels of the dominant colours, and thus three sets of DCT coefficients are obtained. Finally, a few low-frequency coefficients are selected using zigzag scanning and quantised to form the descriptor. Based on previous experimental results [10], 6 coefficients are used for the Y channel and 3 for each of the Cb and Cr channels.

For distance measuring, given two colour layout descriptors $\{Y, Cr, Cb\}$ and $\{Y', Cr', Cb'\}$, the following distance formula is employed:

$$\sqrt{\sum_{i\in(Y)} W_{yi}(Y_i - Y_i')^2} + \sqrt{\sum_{i\in(Cr)} W_{ri}(Cr_i - Cr_i')^2} + \sqrt{\sum_{i\in(Cb)} W_{bi}(Cb_i - Cb_i')^2}, \tag{1}$$

where (Y_i, Cr_i, Cb_i) represent the ith DCT coefficients of the respective colour channels, and (W_{yi}, W_{ri}, W_{bi}) represent the weights of the ith DCT coefficients respectively.

5.2 The Edge Histogram Descriptor

The edge histogram descriptor [11] is an 80-bin histogram representing the distribution of the five types of edges in an image. The five types of edges are the vertical, the horizontal, the 45-degree, the 135-degree and the non-directional. This descriptor can be used for retrieving images of non-homogeneous textures, e.g., natural pictures.

To extract an edge histogram descriptor, the image is first partitioned into 4×4 sub-images, denoted by $I_{(0,0)}$, $I_{(0,1)}$, ..., $I_{(3,3)}$. Occurrence of each of the

five types of edges in each of the $I_{(i,j)}(i,j = 0..3)$ contributes to the value of an appropriate bin of the histogram. Thus five times sixteen gives the number eighty.

To detect edges, each of the sub-images $I_{(i,j)}(i,j = 0..3)$ is then divided into a fixed number of image blocks. (The fixed number is 1100 in our application.) Each of these image blocks is further sub-divided into 2×2 blocks. An edge detector is then applied to each of the image blocks, treating them as 2×2 pixel images. With the edge detector, for an image block, five edge strengths, one for each of the five types of edges, are computed. If the maximum of these edge strengths exceeds a certain preset threshed, the corresponding image block is marked as an edge block, which contributes to the appropriate bin of the descriptor. (A more detailed representation of extracting this descriptor was presented in [11].)

To measure the distance between two quantised histograms $\{h_i\}$ and $\{h'_i\}(i = 0..79)$, with 3 bits representing each h_i/h'_i, we made use of a matrix of quantisation, denoted by QM. (More information about this matrix/table can be found in [11].) This QM has 5 rows and 8 columns, corresponding to 5 types of edges and 8 patterns that 3 bits can represent.

The distance measuring formula employed in our system is

$$d_{(h_i,h'_i)} = \sum_{i=0}^{79} |QM_{(i\%5,h_i)} - QM_{(i\%5,h'_i)}| \; + \; 5\sum_{i=0}^{4} |h_i - h'_i| \; + \; \sum_{i=5}^{79} |h_i - h'_i|, \quad (2)$$

where the % sign denotes the operation of taking the division remainder.

5.3 The Colour Auto Correlogram Descriptor

The colour auto correlogram is a simplified version of the colour correlogram [12]. The colour correlogram descriptor is defined on the HSV colour space, characterising the spatial correlation of pairs of colours in an image.

To formalise this descriptor, let p_1 and p_2 be two pixels in an image I, and (x_1, y_1) and (x_2, y_2) be their coordinates respectively. The distance between p_1 and p_2 is defined by $\max\{|x_1 - x_2|, |y_1 - y_2|\}$, thus, $|p_1p_2| = \max\{|x_1 - x_2|, |y_1 - y_2|\}$. The colours in I are quantised into $c_1, c_2, \ldots, c_n$, and I_{c_i} denotes the set of pixels whose colours are c_i, $i \in \{1, 2, \ldots, n\}$.

The colour correlogram of I is a table indexed by pairs of colours (c_i, c_j), $i, j \in \{1, 2, \ldots, n\}$. For each (c_i, c_j) the d-th entry of the table specifies the probability of finding a pixel of colour c_j at a distance d from a pixel of colour c_i in the image I. (See Formula 3 below.)

$$\gamma^{(d)}_{c_i,c_j}(I) = \Pr_{p_1 \in I_{c_i}, p_2 \in I}[p_2 \in I_{c_j} \,|\, |p_1p_2| = d]. \quad (3)$$

In practice, we often place an upper bound on the considered distances d. Let $d_{\max}$ denote this upper bound. If we consider all the possible combinations, the table will be very large, at the size of $O(n^2 d_{\max})$. Therefore the auto correlogram

is a simplified version in that only pairs of identical colours are considered. Thus we have,

$$\gamma_{c_i}^{(d)}(I) = \Pr_{p_1 \in I_{c_i}, p_2 \in I}[p_2 \in I_{c_i} \mid |p_1 p_2| = d]. \tag{4}$$

Given two images I and I', to measure the distance between their auto correlograms, in our system, the L_1 distance was calculated by the formula

$$d_{(I,I')} = \sum_{i=1}^{n} \sum_{d=1}^{d_{\max}} |\gamma_{c_i}^{(d)}(I) - \gamma_{c_i}^{(d)}(I')|, \tag{5}$$

where n is the number of the quantised colours.

5.4 The Lucene Index

The index creator extracts the three visual descriptors from the images one by one in the collection. After each image has been processed, the three descriptors, together with the canonical path (both absolute and unique) of the image, are saved into a Lucene index.

The Lucene index contains a sequence of *document*s, with each document storing the information of an image. Each document contains four *field*s, corresponding to the three visual descriptors plus the image's canonical path. Each field contains a named pair of *term*s, with the first naming the type of the field, and the remaining ones being the values, either one of the three visual descriptors, or the canonical path.

5.5 Similarity Measurement and Returning Results

When launching a content-based search, the user has to supply through the web interface a sample image I_s together with the following parameters:

- a weight factor w_l for the colour layout descriptor,
- a weight factor w_e for the edge histogram descriptor,
- a weight factor w_a for the auto correlogram descriptor,
- and, the upper bound of the number of the images that will be returned as results, denoted by $r_{\max}$.

The weight factors must satisfy the two conditions.

$$0 \leq w_l, w_e, w_a \leq 1, w_l + w_e + w_a > 0. \tag{6}$$

Once the user has submitted the request, the index searcher will search through the index. For each indexed image I_i, $i \in \{1, 2, \ldots, n\}$, n being the total number of the images, a relevance factor $\alpha_{(I_s,I_i)}$ will be calculated. This $\alpha_{(I_s,I_i)}$ measures the visual distance between I_s and I_i.

Let d_l, d_e, d_a be the distances calculated according to the formulas (1), (2), (5), respectively, and $n_{(w>0)}$ be the number of non-zero weight factors. The $\alpha_{(I_s,I_i)}$ is calculated by the following formula:

$$\alpha_{(I_s,I_i)} = \frac{d_l w_l + d_e w_e + d_a w_a}{n_{(w>0)}}. \tag{7}$$

The relevance factor is guaranteed to fall into the range $[0, 1]$.

The image collection is then sorted by the relevance factors, and the $r_{\texttt{max}}$ most relevant images are returned as the result set for the CBIR part. This result set is then intersectioned with the result set from the keyword-based retrieval. The intersection is returned as the final result to the user.

6 Evaluation

We have performed the tests to demonstrate the advantage that the hybrid retrievals had over those of pure content-based. The results are evaluated by the traditional *precision-recall* measures, in which recall is defined as the fraction of the relevant items that has been retrieved, and precision is the fraction of the retrieved items that are relevant.

To run the tests, we have fixed $r_{\texttt{max}}$ to 20, and the three weights all to 0.5. Five queries were made, in which the pure content-based retrievals use sample images, and the hybrid ones use the same sample images plus keywords. An image in the collection is considered relevant to a sample image if both their semantic meanings and visual effects are relevant. Retrieved images with relevance factors less than 0.3 are not counted as in the result set. The results of the tests are summarised in Table 1. The results have demonstrated that the hybrid retrievals had higher precisions than the pure content-based retrievals at same recall levels.

Table 1. Results of the tests between the pure content-based retrievals and the hybrid ones (Con. stands for pure content-based retrieval. Hyb. stands for hybrid retrieval.)

	Query 1		Query 2		Query 3		Query 4		Query 5	
	Con.	Hyb.	Con.	Hyb.	Con.	Hyb.	Con.	Hyb.	Con.	Hyb.
Relevant in collection	5	5	5	5	5	5	6	6	6	6
Result set	11	7	2	2	11	6	3	2	7	2
Relevant in result set	2	2	1	1	2	2	2	2	1	1
Precision	0.18	0.29	0.5	0.5	0.18	0.33	0.67	1	0.14	0.5
Recall	0.4	0.4	0.2	0.2	0.4	0.4	0.33	0.33	0.16	0.16

7 Conclusion and Future Work

The hybrid image retrieval system presented in this paper indexes and retrieves images based on both the keyword annotations and the visual contents. The purpose of developing such a system was to try to overcome the drawbacks of the pure content-based approaches where the low-level visual descriptors carry no high-level semantic meanings. However, we are fully aware that our work is far from mature. The system currently indexes the text annotations and visual

features separately, and therefore separately they are searched through. In the future we may explore suitable data structures so that both text annotations and visual features may be indexed at the same time and saved in the same index. Thus the need for a database management system may be stripped off.

Acknowledgement. This work was funded by the National Key Technology R&D Program of MOST. The project number is 2006BAK31B03.

References

1. Rui, Y., Huang, T.S., Chang, S.F.: Image Retrieval: Past, Present, And Future. In: International Symposium on Multimedia Information Processing (1997)
2. ISO/IEC/JTC1/SC29/WG11: CD 15938-3 MPEG-7 Multimedia Content Description Interface - Part 3. In: MPEG Document W3703 (2000)
3. Finlayson, G.D.: Color in Perspective. IEEE Transactions on Pattern Analysis and Machine Intelligence 18, 1034–1038 (1996)
4. Huang, T., Mehrotra, S., Ramchandran, K.: Multimedia Analysis And Retrieval System (MARS) Project. In: 33rd Annual Clinic on Library Application of Data Processing - Digital Image Access and Retrieval (1996)
5. Manjunath, B.S., Ohm, J.R., Vasudevan, V.V., Yamada, A.: Colour And Texture Descriptors. IEEE Transactions on circuits and systems for video technology 11(6) (2001)
6. Manjunathi, B., Ma, W.: Texture Features for Browsing And Retrieval of Image Data. IEEE Transactions on pattern analysis and machine intelligence 18(8) (1996)
7. Lux, M.: Caliph & Emir - MPEG-7 Based Java Prototypes for Digital Photo And Image Annotation And Retrieval, Project home page is found via (2007), `http://www.semanticmetadata.net/features/`
8. Busch, M., Cohen, D., Cutting, D.: Apache Lucene - A High-performance, Full-featured Text Search Engine Library in Java, Project home page is found via (2008), `http://lucene.apache.org/java/docs/index.html`
9. Zhou, X.S., Huang, T.S.: Unifying Keywords And Visual Contents in Image Retrieval. IEEE Multimedia 9(2), 23–33 (2002)
10. Kasutani, E., Yamada, A.: The MPEG-7 Color Layout Descriptor: A Compact Image Feature Description for High-speed Image/Video Segment Retrieval. In: IEEE International Conference on Image Processing, pp. 674–677 (2001)
11. Won, C.S., Park, D.K., Park, S.J.: Efficient Use of MPEG-7 Edge Histogram Descriptor. ETRI (Electronics and Telecommunications Research Institute) Journal 24(1) (2002)
12. Huang, J., Kumar, S.R., Mitra, M., Zhu, W.J., Zabih, R.: Image Indexing Using Color Correlograms. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 762–768 (1997)

Use of Artificial Neural Networks in Near-Infrared Spectroscopy Calibrations for Predicting Glucose Concentration in Urine

Weiling Liu[1], Weidong Yang[1], Libing Liu[1], and Qilian Yu[2]

[1] Hebei University of Technology, Tianjin, 300130, China
[2] Tianjin University, Tianjin, 300072, China
goodtimeuk@gmail.com

Abstract. In this paper, Glucose concentration in urine sample were determined by means of Near-infrared spectroscopy combination with artificial neural networks (ANNs). Distinct ANN models were developed for the multicomponent analysis of biological samples. The developed prediction algorithms were tested on unseen data and exhibited high level of accuracy, comparable to the partial least squares (PLS) regression method. Near-infrared (NIR) spectra in the 10000~4000 cm^{-1} region were measured for human urine glucose with various concentrations at 37 °C. The most informative spectral ranges of 4400~4700 cm^{-1} was selected by principle component analysis (PCA) for glucose in the mixtures. For glucose, the correlation coefficient (R) of 0.999 and the root mean square error of prediction (RMSEP) of 20.61 mg/dl were obtained. The presented algorithms can be used as a rapid and reliable tool for quantitative analysis of glucose in biological sample, since the concentration can be determined based on NIR measurement.

Keywords: Near-infrared spectroscopy, Glucose, Artificial neural network, Partial least-squares , Principal component analysis, Urine.

1 Introduction

Urine contains a wide variety of substances. The change of components in urine can not only reflect pathological change in urinary system and reproductive system, but also relate to a lot of whole body diseases. The urine components determined by near infrared spectroscopy (NIR) spectroscopy are glucose, human serum albumin (HSA), total protein, globulin, urea, etc. Glucose can be obtained from our diet as well as from our body tissues. For a normal person, glucose in urine is low. The quantitative determination of glucose is very important to diagnose diseases because the concentration of each component in urine varies with the conditions of health.

Glucose or other components present in urine have distinct spectral fingerprints in the NIR region because of the relatively strong absorption of overtones and the combination modes relative to several functional groups, such as C–H (aliphatic), C–H (aromatic), C–O (carboxyl), O–H (hydroxyl) and N–H (amine and amide), usually present in the organic compounds [5].

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1040 – 1046, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Multivariate calibration models have come into wide use in quantitative spectroscopic analyses due to their ability to overcome deviations from the Beer-Lambert law caused by effects such as overlapping spectral bands and interactions between components. PLS and PCR are the most widely used chemometric techniques for quantitative analysis of complex multicomponent mixtures. These methods are not optimal when the relationship between the IR absorbances and the constituent concentration deviates from linearity[8]. When nonlinear spectral relationships are present in multicomponent mixtures, utilizing principal component scores of the spectra in conjunction with ANN yields better predictions of analytes than those obtained with PLS or PCR techniques alone. The PLS technique produces orthogonal scores similar to those for PCR. In addition, typically, the predictions obtained with PLS require fewer scores compared with those for PCR and thus provide a faster and more parsimonious solution.

In several recent works, artificial neural networks (ANN) were employed as the modeling tool for the IR measurement. Yang et al. [1] employed mid infrared spectroscopy (IR) and artificial neural networks (ANNs) predict the distillation curve points and certain cold properties of diesel fuel fractions. T. Kuzniz et al.[2] applied network training provides quantitative analysis of many pollutants' absorption and fluorescence spectra. The theory and the application of ANNs in modeling chemical data have been widely presented in the literature [10].

In this current work, we have developed ANN models for the prediction of glucose concentration in urine.

2 Model Definition

The same model is applied to absorption measurements and is based upon the well-known Beer Lambert law . For low glucose concentrations we have:

$$A(\{c_i\},\lambda) = \sum_i c_i \cdot A_i(\lambda) \ , \tag{1}$$

where $A_i(\lambda)$ is the optical density at wavelength λ, of the *i*th component, whose concentration is c_i, $A(\lambda)$ the optical density of the mixture. If we have the spectra of all the components $\{A_i(\lambda)\}$ and measure the spectrum of the mixture $\{A_i(\lambda)\}$, we can use a mathematical tool to estimate the $\{c_i\}$.

The spectrum of the urine sample is a plot of the yield Y(E), which contains glucose concentration information. In terms of mathematical abstract, the problem of spectral data analysis can be viewed as a pair of forward and inverse transformation between the input vector – the spectra data and the output vector – the structural parameters of a sample. We use principal component analysis (PCA) to reduce the dimensionality of the input data. In the present application, we performed a principal component analysis for the dataset containing a few hundred spectra of glucose absoroption. As the result, it was found that retaining 8 principal components can account for 99.6% of variation in the dataset.

We built a multilayer feed forward perception network (MLP) to analyze our dataset, where the structural information of a sample could be predicted from the corresponding spectra available. Initially a number of different network architectures were tested. After some trials, three layer networks were found to be appropriate. The network consists of three layers of nodes: one input layer, one output layer and a hidden layer. All the nodes in one layer are connected to the next layer feed forward, as shown in Fig.1. The input is the vectors of principal components. The sample structural parameters are output variables at the nodes of output layer. The number of nodes at the hidden layer is adjustable so that an optimal architecture can be achieved through experiments.

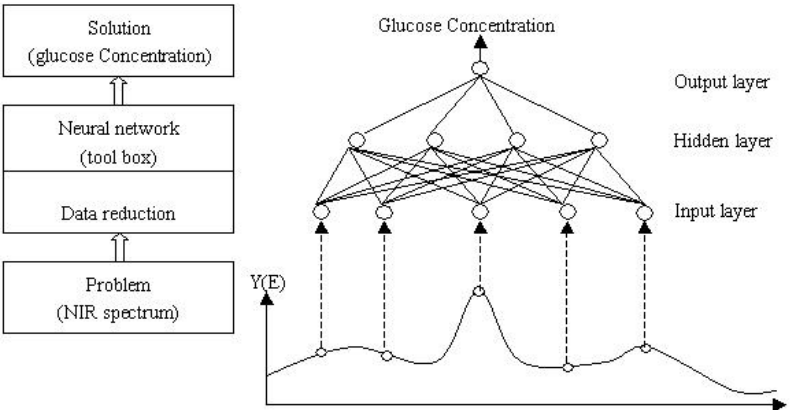


Fig. 1. Schematic diagram of neural networks for solving spectral analysis problem. Values of spectra at energy positions E1, E2,…Em are as input for the network. Structural parameters are the output.

3 Experiment Section

3.1 Instrumentation

The NIR transmittance spectra in the 10000–4000 cm^{-1} region were measured by using a Perkin Elmer Spectrum GX FT-NIR spectrometer equipped with an Indium Galium Arsenide (InGaAs) detector. The settings of the spectrometer where optimized for NIR measurement by using its 100W tungsten-halogen lamp, its InGaAs thermoelectrically cooled detector, and its quartz beam splitter. The cuvette was a quartz flow through cell with a pathlength of 1mm. To achieve a satisfactory signal-to-noise ratio, 32 spectra were averaged at a spectral resolution of 2 cm^{-1}. All the spectral data were collected with a 4 cm^{-1} spectral resolution, and 256 scans were co-added to ensure a sufficient signal-to-noise ratio. A quartz cell with a path length of 1 mm was used throughout the experiments and this empty cell was measured as background. The cell was cleaned 3–4 times by the distilled water and then 2–3 times by a new sample solution before each measurement. The temperature of sample was kept at 37 ± 0.2°C by using a temperature controller that was connected to a cuvette cell

holder with fiber optics. Sample was pumped through tube with a peristaltic pump such that the sample would arrive at the cuvette at the desired temperature.

3.2 Software

The SPECTRUM (ver.3.01: Perkin Elmer, USA) software program was employed for the spectral data collection. The spectral data obtained were converted into files for the Unscrambler (ver.7.6: CAMO ASA, OSLO, Norway) software program. The programs used for data pre-processing and for the PCA or PLS calculations were Unscrambler. The ANN calculations were carried out by a laboratory-written program in MATLAB for Windows (Mathworks, USA, Version 6.5).

4 Results and Discussion

4.1 Development of the ANN Models

In order to predict urine glucose concentration from recorded spectra multivariate calibration, models were developed by artificial neural networks (ANN). The NIR spectra of urine sample were shown in Fig. 2. Since ANN are universal approximators capable to fit any continuous linear and non-linear relationship (function), existing between dependent and independent variables.

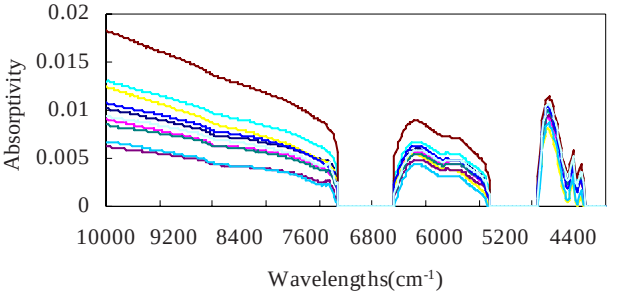


Fig. 2. NIR spectra of ten urine samples

The independent variables to be modeled are presented to the input layer of the ANN and then weighted by wij between the input and the first hidden layer. Hidden layer nodes receive the weighted signals and after summation, they transform them using a non-linear sigmoid function S(.). The output of each node from the 1st hidden layer, after the weighting, enters the 2nd hidden layer where a new summation and transformation with the same non-linear function takes place. Finally, the output of the hidden layers, weighted with the respective layer weight, enters the output layer, where it is transformed with a linear function L(.) to the final output of the model. Mathematically, an ANN model can be written as: $y = L\left(w_3 S\left(w_2 S\left(w_1 x\right)\right)\right)$ where x denotes the input data vector, w1, w2, w3 denote the weights of the successive layers and y is the dependent variable of the data set. For the development of the ANNs, the

data set was randomly split into two subsets. Each network was trained to the desired accuracy level using the first (training) set containing 85% of the samples and their performance was evaluated using the second (test) set, i.e. on unseen by the network data.

The principal component analysis (PCA) technique was applied to the derivatives data set, in order to reduce its dimensionality, since the number of independent variables (absorbances at 500 wavenumbers) was significantly greater than the number of the available samples. PCA derives a new orthogonal set of variables, called the principal components (PCs) that are linear combinations of the original variables and attempt to account for the largest possible portion of the total variance in the data set. Each principal component defines a different direction of variance contained in the original data set and therefore can be considered as a new 'derived' property of the samples. The orthogonality constraint, imposed by the mathematics of the PCA, ensures that these 'new' variables are uncorrelated. The PCA was performed using the Unscrambler software. The analysis of the input matrix showed six significant eigenvalues, with the first eight PCs describing 98.7% of the original variance. The first principal component versus the second principal component was plot in Fig. 3, PC1and PC2 represent about 93% of data variance. The first eight PCs were used unscaled as the input data set for the development of the ANNs.

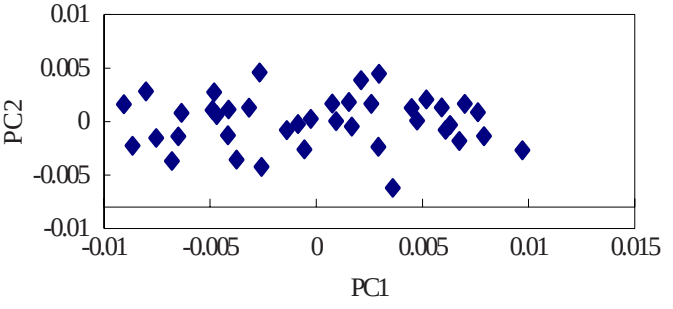


Fig. 3. Plot of the first principal component scores vs. the second principal component scores

4.2 Calibration and Prediction Results

The performance of the different models was evaluated in terms of root mean square error of prediction (RMSEP) for the test samples, given by Eq. (2):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2}{N}} \,. \tag{2}$$

The root mean square error of prediction (RMSEP) is an indication of the average error in the analysis, for each component. y_i is the true concentration of the analyte in sample i, $\hat{y}_i$ represents the estimated concentration of the analyte in sample i, and N is the total number of samples used in the prediction set.

In the ANN modeling, the input variables were the first eight scores of the PCA analysis because with this number 98.7% of data variance can be explained. Addition of more principal components only adds noise into the model. The number of hidden units (centers) selected was 13 and for this model, the calculated root mean square error of prediction for the test data set was 0.25%. This approach resulted in the best model, with a correlation coefficient equal to 0.96.

Fig. 4 shows the predicted glucose concentrations in urine samples from the validation set obtained with the PLS model (left side)and the ANN model (right side). The ANN approach with the use of principal component scores as preprocessed inputs has been shown to provide better predictions of analyte concentrations than PLS in the presence of nonlinear spectral relationships.

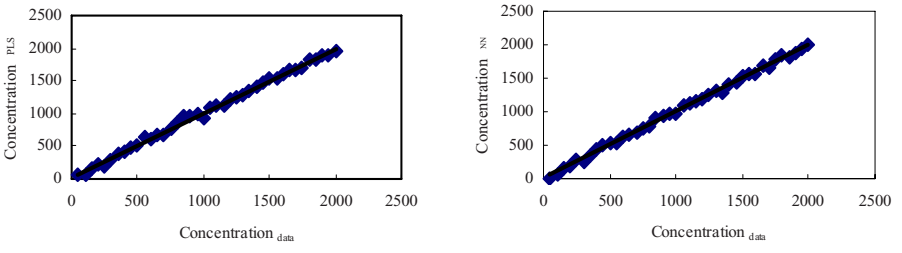


Fig. 4. Comparison of PLS (RMSEP: 37.5mg/dL) and ANN (RMSEP: 26.1mg/dL) calibration curves for the NIR determination of glucose in urine

5 Conclusion

Urine is an extremely difficult and complex biological system due to the high spectral overlapping observed between the absorption of glucose and albumin.

Artificial neural networks were used to predict glucose concentration. The developed models predict accurately these properties based on the FT-IR signal is better than PLS methods. These results indicate that urine sample spectroscopy in the near infrared region with data treatment by ANN, is faster, clean and reliable. This procedure may be used as a tool for determination of other materials in biological samples.

References

1. Pasadakis, N., Sourligas, S., Foteinopoulos, C.: Prediction Of The Distillation Profile and Cold Properties of Diesel Fuels Using Mid-IR Spectroscopy and Neural Networks. J. Fuel. 85, 1131–1137 (2006)
2. Kuzniz, T., Halot, D., Mignani, A.G., Ciaccheri, L., Kalli, K., Tur, M., Othonos, A., Christofides, C., Jackson, D.A.: Instrumentation for the Monitoring of Toxic Pollutants in Water Resources by Means Of Neural Network Analysis of Absorption and Fluorescence Spectra. J. Sens. Actuators B: Chem. 121, 231–237 (2007)
3. Michael, M., Li., F.X., Kevin, T.: Principal Component Analysis and Neural Networks for Analysis of Complex Spectral Data from Ion Backscattering. In: Proceedings of the 24th IASTED International Multi-Conference, pp. 228–234. ACTA Press, USA (2006)

4. Kuzniz, T., Tur, M., Mignani, A.G., Romolini, A., Mencaglia, A.: Neural Network Analysis of Absorption Spectra and Optical Fiber Instrumentation for the Monitoring of Toxic Pollutants in Aqueous Solutions. In: Proceedings of SPIE, vol. 4185, pp. 444–447. Society of Photo-Optical Instrumentation Engineers, Bellingham (2000)
5. Fidêncio, P.H., Poppi, R.J., de Andrade, C., João,: Determination of Organic Matter in Soils Using Radial Basis Function Networks and Near Infrared Spectroscopy. J. Analytica Chimica Acta. 453, 125–134 (2002)
6. Mendelson, Y., Clermont, A.C.: Blood Glucose Measurement by Multiple Attenuated Total Reflection and Infrared Absorption Spectroscopy. J. IEEE Trans. Biomed. Eng. 37, 458–465 (1990)
7. Goicoechea, H.C., Olivieri, A.C.: Determination of Theophylline in Blood Serum by UV Spectroscopy and Partial Least Squares (PLS-1) Calibration. J. Analytica Chimica Acta. 384, 95–103 (1999)
8. Bhandare, P., Mendelson, Y., Peura, R.A.: Multivariate Determination of Glucose in Whole Blood Using Partial Least-Squares and Artificial Neural Networks Based on Mid-Infrared Spectroscopy. J. Applied Spectroscopy 47, 1214–1221 (1993)
9. Smits, J.R.M., Malssen, W.J., Buydens, L.M.C., Kateman, G.: Using Artificial Neural Networks for Solving Chemical Problems. Part 1. Multi-layer Feed-forward Networks. J.Chemometrics and Intelligent Laboratory Systems 22, 165–189 (1994)

A Rotation Angle Calculation Method about Bearing Images Based on Edge Image Correlations

Lin Liu

Infromation & Electronics Engineering college of Zhejiang Gongshang University Hangzhou, Zhejiang, China, 310018
LiuLin@zjgsu.edu.cn

Abstract. In Industrial automation field rotation angle calculation between successive bearing images is frequently required and the rotation angle is used for locating the oil feeder. Because bearing images are rotational symmetric, also they have perspective distortion, the phase correlation method in frequency domain usually cannot find reasonable rotation angle. This paper presents a new method based on edge images' correlations, at first original images' Canny edges are got after polar coordinate transform, then Hamming distance is used to calculate spatial correlation between bearing images, at last reasonable rotation angles are got when Hamming distance get the minimum value. For this kind of rotational symmetric images with rich edges, experiments show that the proposed method can efficiently solve this problem with much low computation cost.

Keywords: Rotation angle, rotational symmetric property, polar transform, edge images, image registration, phase correlation.

1 Introduction

One important step of bearing automatic production is oil filling on bearings balls for finished bearings, in order to lubricate bearings and let bearing be rust preventive. Oil filling is performed by oil feeder on operation desk, which is almost the same size as bearing. Every time rotation angles of bearings balls on operation desk are different, so oil feeder need to be rotated for accurate location. All these relocation operations are implemented with the help of computer vision and image treatment.

Two successive bearing images are showed in figure 1 and Figure 2. They are shot by back light and direct light respectively and their size in images are 4 times of the size of an object. There are 7 balls and 7 ball joints between inner ball collar and outer ball collar. In Fig.1 ball joints occupy dim parts and balls occupy bright parts in Fig.1, but in Fig.2 ball joints occupy white parts and balls occupy black parts. Among these two groups of images, distances between balls are fixed, but corresponding balls in successive images are located in different position, and they produce a rotation angle which is exactly what we want.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1047 – 1055, 2008.
© Springer-Verlag Berlin Heidelberg 2008

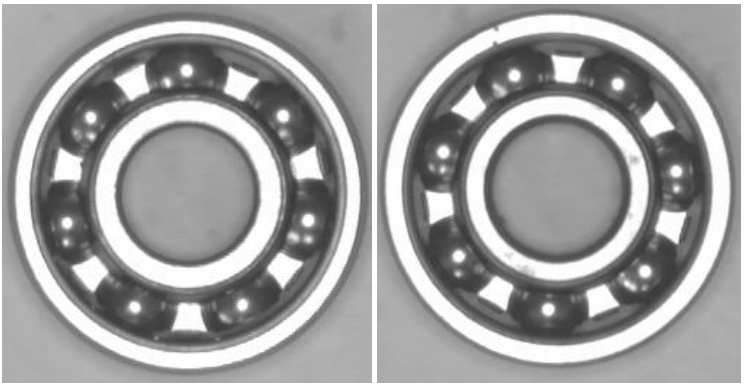


Fig. 1. Two successive bearing images shot by direct light

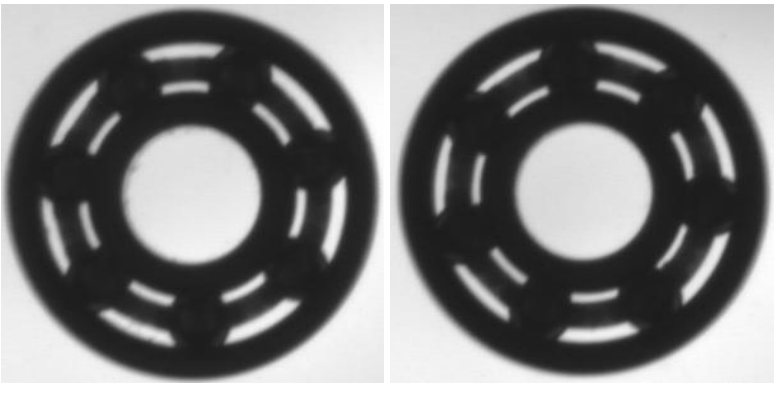


Fig. 2. Two successive bearing images shot by back light

2 Current Methods for Image Rotation Angle Calculation

In fact rotation angle calculations in figure 1 and figure 2 are image registration between these two images. For image registration, one kind of method is based on Hough transform[1-3], but it cost much computation; another prevailing method based on phase correlation in frequency domain is widely used[4-7], which cost less computation and it is robust against image noise. Phase correlation method is initially used for image registration with translation motion, later it is also used for image registration with image rotation and image scale. If image translation and scale are not considered, image rotation angle calculation by phase correlation method can be described as follows.

If an image $f_2(x, y)$ was obtained by the image rotation of original image $f_1(x, y)$, and the rotation angle is θ_0, that is:

$$f_2(x, y) = f_1(x\cos\theta_0 + y\sin\theta_0, -x\sin\theta_0 + y\cos\theta_0) \tag{1}$$

If the Fourier transform of $f_1(x,y)$ and $f_2(x,y)$ are $F_1(\varepsilon,\eta)$ and $F_2(\varepsilon,\eta)$ respectively, so according to properties of Fourier transform:

$$F_2(\varepsilon,\eta) = F_1(\varepsilon\cos\theta_0 + \eta\sin\theta_0, -\varepsilon\sin\theta_0 + \eta\cos\theta_0) \tag{2}$$

If $M_1(\varepsilon,\eta)$ and $M_2(\varepsilon,\eta)$ are amplitudes of $F_1(\varepsilon,\eta)$ and $F_2(\varepsilon,\eta)$ respectively, then:

$$M_2(\varepsilon,\eta) = M_1(\varepsilon\cos\theta_0 + \eta\sin\theta_0, -\varepsilon\sin\theta_0 + \eta\cos\theta_0) \tag{3}$$

Obviously $M_1(\varepsilon,\eta)$ and $M_2(\varepsilon,\eta)$ have the same amplitudes, only differs by a rotation transform. Because it is difficult to find the value of θ_0 directly, so a polar transform is performed on equation (3), which changes rotation transform to translation transform:

$$M_2(\rho,\theta) = M_1(\rho,\theta-\theta_0) \tag{4}$$

Here $\rho = \sqrt{\varepsilon^2+\eta^2}$, $\theta = \tan^{-1}(\eta/\varepsilon)$. Fourier transform was made on amplitude of $M_1(\rho,\theta)$ and $M_2(\rho,\theta)$, and we set $X_1(\rho,\theta) = |M_1(\rho,\theta)|$, $X_2(\rho,\theta) = |M_2(\rho,\theta)|$. According to translation invariance to Fourier transform, it can be found that $X_1(\rho,\theta)$ and $X_2(\rho,\theta)$ have the same amplitude, but with a θ_0 phase difference. Inverse Fourier transform is performed on cross power spectrum of $X_1(\rho,\theta)$ and $X_2(\rho,\theta)$, then a maximum value of the inverse Fourier transform can be found at coordinates (0, θ_0)[8,9].

After phase correlation calculation, the result is showed in figure 4. Ideally there is only one extreme point in figure 4 whose coordinates correspond to the rotation angle, but there exists several extreme points in figure 4 which distributes on two lines of 0 and 360 (which are produced by the periodicity of Fourier transform). The maximum point coordinates at point (0,180), which means the rotation angle is 180 degrees, obviously it is not a reasonable result according to figure 1.

The failure reason lays on two points, one point is that while two successive bearings are the same, two bearings' positions on operation desk have imperceptible

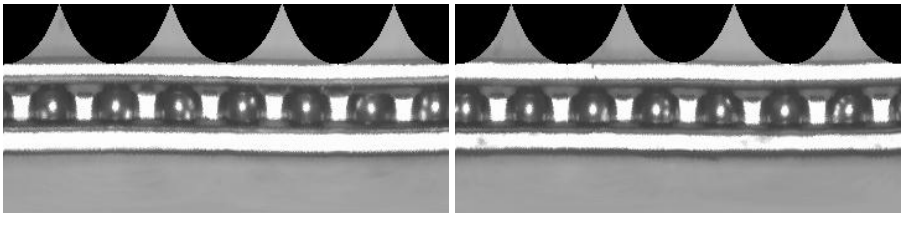


Fig. 3. Polar transform result of the images in Fig.1

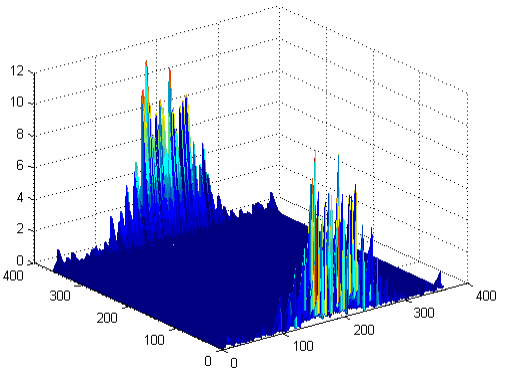


Fig. 4. The rotation angle calculation result of the images in Fig.1 base on phase correlation method

displacement, which causes the perspective distortion during shot and produces two images with minor difference. Because the bearing images are magnified by original objects, these differences cannot be ignored. Another point is that because bearing images are rotational symmetric, global optimal solution is not reasonable, and rotation angles must be restricted as a certain range (here $0 \le \theta_0 < 360/7$ degree).

3 A Method Based on Edge Image Correlations

This paper presents a method based on edge images correlation. Here edge image is defined as binary image after edge enhancement. Edge images is taken because the following facts: rotation angle calculation is essentially related with the shape of bearing image and has little to do with gray level of bearing image, and image edges can roughly represent image's shape. So here we calculate rotation angle by edge images correlation, which can make full use of high frequency parts of original images, also it can reduce the disturbance of low frequency parts of original images.

Similar with the phase correlation method, here edge images and their correlation are also acquired after original images' polar transform, in another word the calculation is implemented on polar coordinate images. Because edge images are binary images, their correlation can be directly calculated by Hamming distance. Here the difference of two successive images $f(x, y)$ and $g(x, y)$ in polar coordinate can be defined as follows:

$$D_{fg}(r) = \sum_{x=1}^{M} \sum_{y=1}^{N} \left| f(x, y) - g^{r}(x, y) \right| \quad (-MA \le r \le MA) \tag{5}$$

Here $D_{fg}(r)$ represents the difference of two images with rotation angle r, when $D_{fg}(r)$ get the minimum value r represent the wanted rotation angle. M and

N represent the width and height of two images respectively. $f(x,y)$ represents the pixel value of the preceding image at coordinate (x,y), its value is 0 or 255. g^r represents a new image obtained by cyclical right shift r pixels of the latter image, and $g^r(x,y)$ represents the pixel value of new image at coordinates (x,y). The value of r ranges in $[-MA, MA]$, and it represents the range of rotation angle; in the case of figure 1 and figure 2, $MA = 360 \times N/7$ represents the circle is divided by 7 parts and the rotation angle accurate to $1/N$ degree, where N is a natural number.

$$\begin{cases} g_{ij}^{r} = g_{kj}^{0} \\ k = (i+r) \bmod M \end{cases} \tag{6}$$

Image cyclical right shift is implemented on polar coordinate image according to equation (6), here M represents the image width. The result of image cyclical right shift is that pixels move out from right edge appear on left edge of the new image, which corresponds to the rotational symmetric property of figure 1 and figure 2.

Except all above process, image preprocessing is performed to ensure the reliability of rotation angle calculation. At first every original gray image is binarized, then the binary image is connected-area labeled to get accurate bearing image border. Before further treatment, the similarity of two images are calculated by first order central moment[10], in order to ensure two bearings belong to the same kind and images are shot by the same direction light(that is two images are all back shot or all direct shot), here the similarity value of the same kind must larger than 0.95. At last the image are performed a simple horizontal or vertical correctness to reduce perspective distortion, the correctness equation is as equation (7).

$$\begin{cases} g'(x,y) = g(a \times i, b \times j) \\ a = M_1 / M_2 \\ b = N_1 / N_2 \end{cases} \tag{7}$$

It is hard to get the parameter of affine transform for accurate distortion correctness, here only a simple correctness function is used. After correctness of the latter image according to preceding one, two successive images should have the same size and their perspective distortion on horizontal direction or vertical direction come near. Assume M_1, M_2 are original image width of these two images, and N_1, N_2 are their height respectively. Here $g'(x,y)$ represents the pixel value of corrected image at coordinates (x,y), and its corresponding pixel at the corresponding original image is $g(a \times i, b \times j)$. a, b are correctness ratio as equation (7).

4 Experiments Result and the Analysis

4.1 The System Work Flow

The system flowchart is showed in figure 5, which presents the implement order of all steps above. Image binarization is based on Otsu's method, and 4-connect area labeling is used for searching individual region and getting accurate image border. Edge image is calculated by Canny operator, and performed on gray image. Polar transform is performed on Canny edge image and the origin of polar transform is at the center of the image.

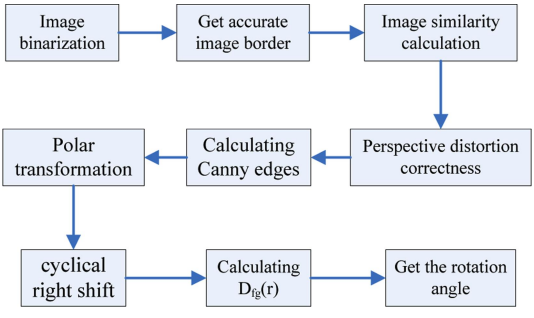


Fig. 5. The system flowchart

4.2 Experiment Result

After polar transform and perspective distortion correctness, the Canny edge images of figure 1 and figure 2 are shows as figure 6 and figure 7 respectively.

The relationship of image difference $D_{fg}(r)$ and rotation angle r in figure 6 and figure 7 is presented in figure 8 and figure 9 respectively. It can be found that in figure 8 $D_{fg}(r)$ get the minimum value at $r = 20$, which implies the rotation angle in figure 1 is 20°; In figure 9 $D_{fg}(r)$ get minimum value at $r = 32$, which implies the rotation angle in figure 2 is 32°.

Fig. 6. Canny edge images of two images in Fig.1

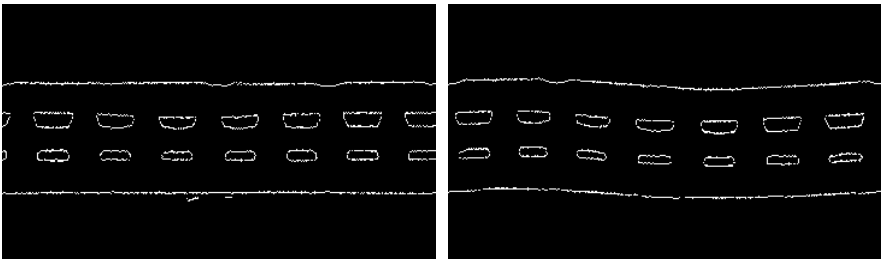


Fig. 7. Canny edge images of two images in Fig. 2

4.3 Computation Complexity Compare

It can be found that the proposed method and phase correlation method both have the computation complexity of $O(N^2)$, and N is total number of image pixels. But correlation computation in the proposed method is integer arithmetic, and it cost much less float arithmetic than phase correlation method, so it has prevailing computing speed.

4.4 Computation Reliability

The computation reliability is listed in table 1. The proposed method was tested by total 111 image pairs shot by direct light or back light. The actual rotation angle is computed by manual, and if the final computation angle varies in $\pm 2^{\circ}$, the result is considered to valid. From table 1 it can be deduced that the proposed method can give a relatively reliable computation result. The correct ratio in back light is higher that in direct light because there exist more detail in direct light image which will affect the final calculation result.

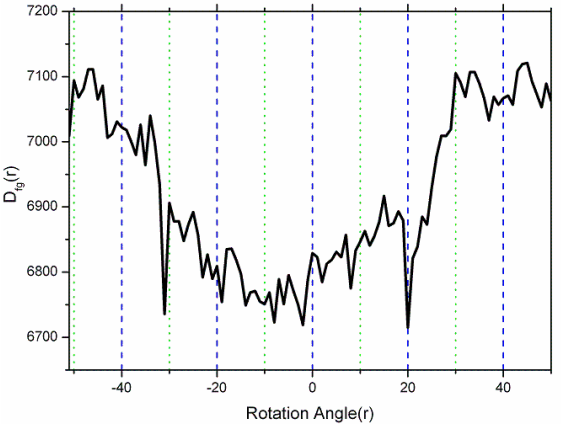


Fig. 8. The relationship of $D_{fg}(r)$ and r in Fig.6

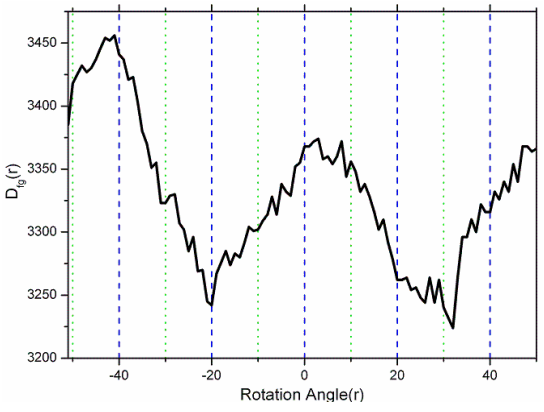


Fig. 9. The relationship of $D_{fg}(r)$ and r in Fig.7

Table 1. Computation reliability

	Direct light	Back light
Total image pairs	56	55
Number of valid result	51	54
Correct rate	91.1%	98.2%

5 Conclusion

In Industrial automation field rotation angles calculation between two successive bearing images is an important part. Because of perspective distortion two successive bearing images are not identical, in addition bearing image is rotational symmetric, so traditional method based on phase correlation cannot find the reasonable rotation angle. In fact shape and other high frequency information decides the rotation angle calculation, and image edges can roughly represent image high frequency information, so this paper gives a rotation angle calculation method based on the correlation of edge images. During preprocessing step, perspective distortion of two successive image is corrected, then their Canny edge images are calculated, next they are transformed to polar coordinates, and one of them is cyclical right shifted, at last correlation based on Hamming distance is calculated and rotation angle appears when the Hamming distance get the minimum value. Because edge images represent shape information and it prevent the disturbance of low frequency information, so compared with phase correlation method this method can get a better reasonable result. At the same time, this method is implemented in spatial domain without Fourier transform, so it can efficiently reduce the floating point arithmetic and improves calculation efficiencies. Experiments also show the effectiveness of this method.

References

1. Onishi, H., Suzuki, H.: Detection of Rotation and Paraller Translation Using Hough and Fourier transforms. In: Proceeding of IEEE International Conference on Image Processing, pp. 827–830 (1996)
2. Pao, D.C.V., Li, H.F., Jayakumar, R.: Shape's Recognition Using the Straight Line Hough Transform: Theory and generalization. IEEE PAMI 14(11), 1076–1089 (1992)
3. Li, Z., Yang, X.H.: A Method of Image Registration for Two Rotated and Translated Images. Journal of Applied Science (China) 23(3), 282–286 (2005)
4. Kuglin, C.D., Hines, D.C.: The Phase Correction Image Alignment Method. In: IEEE International Conference on Cybernetics and Society, New York, USA, pp. 163–165 (1975)
5. Hill, L., Vlachos, T.: On the Estimation of Global Motion Using Phase Correction For Broadcast Application. In: Proceeding of IEEE International Conference on Image Processing, pp. 721–725 (1999)
6. Milanfar, P.: Projection-based Frequency-domain Estimation of Superimposed Translation Motions. Journal of the Optical Society of America A 13(11), 2151–2162 (1996)
7. Zitova, B., Flusser, J.: Image Registration Methods: A Survey. Image and Vision Computing 21(11), 977–1000 (2003)
8. Reddy, B.S., Chatterji, B.N.: An FFT-based Technique for Translation Rotation, and Scale-Invariant Image Registration. IEEE Transaction on Image Processing 5(8), 1266–1271 (1996)
9. Li, Z., Mao, Y., Wang, Z.: A Method of Image Mosaic Using Log-polar Coordinte Mapping. Journal of Image and Graphics(China) 10(1), 59–63 (2005)
10. Belkasim, S.O., Shridhar, M., Ahmadi, M.: Pattern Recognition With Moment Invariants: Comparative Study and New Results. Pattern Recognition 24(12), 1117–1138 (1991)

Image Restoration Using Piecewise Iterative Curve Fitting and Texture Synthesis

Ke Sun, Yingyun Yang, Long Ye, and Qin Zhang

Information Engineering School
Communication University of China, 100024 Beijing, China
ksun@cuc.edu.cn, yingyun6903@163.com,
{yelong,zhangqin}@cuc.edu.cn

Abstract. In this paper, we present a novel scheme for parsing images into medium level vision representation: contour curve and region texture. This scheme is integrated with *piecewise iterative curve fitting* and *texture synthesis*, in which an original image is analyzed at the representation side so as to obtain contours and local texture exemplars. The contour of each region will be processed piecewise by iterative curve-fitting procedure. At the reconstruction side, transmitted coefficients denoted to structural information of regions will be used to reconstruct contours. Texture synthesis algorithm is adopted to restore the rest. Experimental results indicate that our scheme can be done successfully for images with massive structural and stochastic textures. Compared with JPEG, our method can get more coding gain but remain vast information people care.

Keywords: Image restoration, Region extraction, Iterative curve fitting, Texture synthesis.

1 Introduction

The mainstream of image coding scheme is mainly based on hybrid coding framework by which most spatial redundancy, temporal redundancy and statistic redundancy will be eliminated. However, it is difficult to achieve farther prominent advance of compression efficiency since the limitation that pixel-based image representation bring [1]. Then how to find a novel image representation method becomes a hot topic.

Actually speaking, many images contain massive textures like grass, sand, brick, *etc.*, which compose the background. Patrick Ndjiki-Nya *et al*. [2] assume that the textures in an image can be classified into two categories: textures with unimportant subjective details and the remainder. In this paper, we utilize this idea and consider textures as unimportant parts according to the imperfections inherent to the *Human Visual System (HVS)*. These parts can be considered as detail-irrelevant textures if the original image is not shown to the viewer. As a result, changes of texture details will not affect the subjective comprehension toward the original texture. It indicates that vision character can be utilized to remove perceptual redundancy as well as represent images effectively and efficiently.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1056 – 1063, 2008.
© Springer-Verlag Berlin Heidelberg 2008

In addition, there have emerged large quantities of efficient and effective algorithms of texture synthesis in recent years. A nonparametric learning includes many algorithms [3], as practiced by Efros and Leung for texture synthesis [4], and by several follow-up works [5, 6] for super-resolution, most of which tend to preserve the global structure nicely. So it adapts to deal with the structural/deterministic texture as well as the stochastic texture. The synthesized texture is perceived as similar, although it must contain vast variation of the original in pixel level.

Besides the texture, the contour and region shape features have been researched far and wide. Most of the early shape description techniques used binary images. Some examples of binary edge based methods are polygonal approximation, frequency domain representation of curvature, and syntactic description [7]. For polygonal approximation, the boundary of the form is approximated by line segments [8], or in some cases line segments and arcs [9]. In this way, a polynomial function can be used to represent curves and fitted to the contour more effectively.

The inspiration for our work comes from the thought expounded above. Fig. 1 depicts our proposed image representation and reconstruction scheme. In our framework, image segmentation is first performed on the original image. As a result, some regions regarded as the basic cell of image analysis are gained. At the representation side, all regions will be processed one by one. For the contour of each region, a shape describing approach integrated with contour tracing and iterative curve fitting is done to get shape feature. Meanwhile, a small texture sample is selected from the inner texture of each region. At the reconstruction side, contour restoration and texture synthesis are performed successively, both of which utilize the datum attained before to generate the reconstructed image. The flag information is regarded as assistant datum for labeling the serial number of regions.

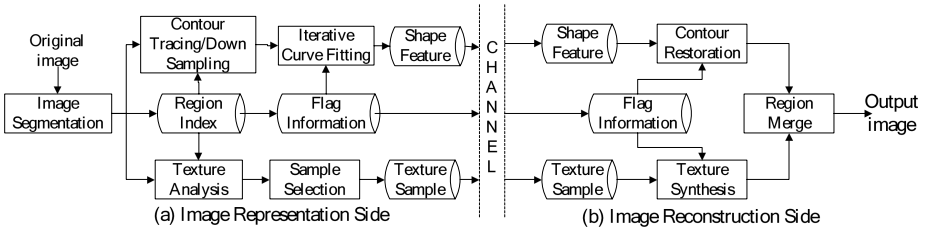


Fig. 1. Our proposed scheme

The remainder of the paper is organized as follows. In Section 2 we describe the feature extraction, while in Section 3 the image restoration is presented. Finally, in Section 4 the experimental results are shown.

2 Image Feature Representation

Shape and texture are salient features of the appearance of objects in natural scenes, and their describing approach will be discussed respectively. Just like what we say in

section 1, in our proposed scheme, we consider the mid-level region as image representing and reconstructing unit which contains shape and texture characteristic.

2.1 Region Extraction

An input image is first segmented into some regions in terms of homogeneous color and texture. The segmentation method we used is the so-called *JSEG* algorithm [10]. Considering that texture sample selection is sensitive to segmentation results, we merge connected small regions if they are nearby in space so that fewer and bigger regions are resulted for each image [11]. After segmenting, homogeneous pixels form one region. It is expedient to get region index list which is used to index each obtained region. Each element of the index list corresponds to the pixel of original image. Every pixel belonged to one region is assigned the same number. This list plays an important role in following processing.

2.2 Representation of Structural Information

After regions are extracted, a region shape describing scheme concludes contour tracking, down sampling and piecewise iterative curve fitting is performed, and Fig.2 indicates the schematic drawing of three stages above.

A. Contour tracing and down sampling. In order to elaborate conveniently, we just take into account one region which should be binarized previously, as is depicted in Fig. 2(a). The algorithm we use has been introduced in [12]. In the first step, an ordered sequence of contour points approximating the input shape is extracted. For this purpose, a contour tracing algorithm (also known as border following or boundary following) similar to that of Pavlidis [13] is applied to the input binary image, showed as Fig. 2(b). The algorithm ignores any "holes" present in the shape providing an ordered sequence of the contour pixels. The tracing algorithm can be easily and efficiently implemented using look-up tables for selecting the pixels in the current pixel neighborhood for checking and deciding in which order they should be checked.

In the next step, a vector of equally distributed points along the curve is extracted from the ordered sequence of contour pixels. It extracts contour points in such a way that they will be approximately equally distributed along the contour, like Fig. 2(c). From this point on, the contour of the region is represented by the ordered sequence (x_s, y_s) of contour points, where s denotes the position along the contour.

B. Piecewise iterative curve fitting. Here, a shape feature representation, we call it piecewise iterative curve fitting, combined with top-down split algorithm and curve-fitting principle is proposed. Attained the boundary of the region, we implement five steps as follows:

1) Take the line segment connecting the two farthest points of the contour, and split the closed contour into two curves. As shown in Fig. 2(d), the blue dashed line segment *AB* partitions the contour into two curve segment, namely curve *AB* and *BA*, because there is a longest distance between point *A* and *B*.

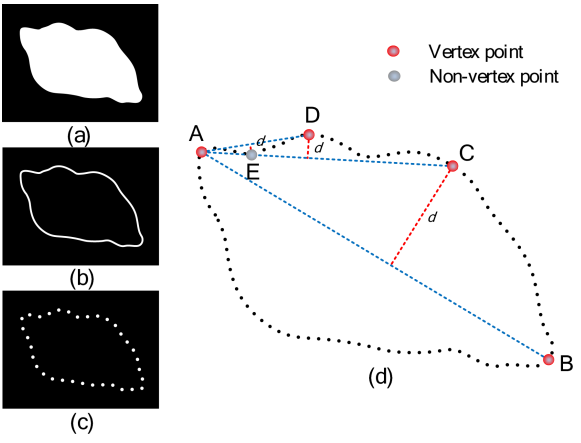


Fig. 2. Schematic drawing of shape describing scheme

2) Connect the starting and ending edge points P_A and P_B into a line segment, then evaluate the distance of each edge point to this line using eq. (2), and determine the edge point P_C farthest from the line segment.

Suppose the coordinates of P_A and P_B are (x_A, y_A) and (x_B, y_B), respectively, so the equation of line passing from the two end points is:

$$x(y_A - y_B) + y(x_A - x_B) + x_A y_B - x_B y_A = 0 \tag{1}$$

The distance d of edge point (x_s, y_s) from the line:

$$d_s = r_s / \Delta_{AB} \tag{2}$$

Where s denotes the position of any point along the curve AB, and

$$r_s = x_s(y_A - y_B) - y_s(x_A - x_B) + x_A y_B - x_B y_A \tag{3}$$

$$\Delta_{AB} = \|A - B\| \tag{4}$$

Thus, the maximum absolute error is:

$$\mathrm{MAE} = \max_{s \in [A,B]} |d_s| \tag{5}$$

Clearly, point C with MAE is the farthest edge point from the line segment AB in Fig. 2(c).

3) If the maximum error of that point from the line segment is above a threshold, split the segment into two segments at that point (i.e. new vertex point C), one between P_A and P_C, the other between P_C and P_B.

4) Repeat steps (2) and (3) separately for the above two sub-segments to determine more and more vertices until the error is smaller than the "split threshold" which has been set previously.
5) Implement polynomial curve fitting on each sub-segment, and the polynomial coefficients are deemed as shape features.

Simply, we only take sub-segment *AD* as an example. The curve *AD* is represented by the order sequence $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ of edge points. A polynomial always may be expressed by the sum of some orthogonal polynomial. Suppose the fitted polynomial is:

$$y = p_0 + p_1 x + p_2 x^2 + \cdots + p_m x^m \tag{6}$$

We may calculate the polynomial coefficients $p_0, p_1, \cdots, p_m$ according to the fundamental curve fitting principle. In this way, a curve is figured by these coefficients and necessary endpoints.

2.3 Texture Exemplar Selection

In the region, a big texture exemplar contains a majority of local and global characteristics of the inner texture. The texture exemplar should be preserved for texture synthesis. Firstly, we analyze the inner texture in a 2D-autocorrelation statistic analysis method [14]. The autocorrelation function of an image can be used to assess the amount of regularity as well as the fineness/coarseness of the texture present in the image. Formally, the autocorrelation function of an image is defined as follows:

$$\rho_{II}(x,y) = \frac{MN}{(M-x)(N-y)} \cdot \frac{\sum_{u=1}^{M-x}\sum_{v=1}^{N-y} I(u,v) I(u+x, v+y)}{\sum_{u=1}^{M}\sum_{v=1}^{N} I^2(u,v)} \tag{7}$$

For regular structural textures, the function will exhibit peaks and valleys. We can easily determine the scale of the texture primitives from the given texture. The exemplar should contain 2~5 texture primitives. For stochastic textures, correspondingly, the size of exemplar may be set smaller.

3 Image Reconstruction

We know that a curve can be described as a polynomial function in eq. (6). Where, the polynomial coefficients $p_0, p_1, \cdots, p_m$ have been attained in section 2.2.*B*. According to the x-coordinates of two endpoints, such as x_m and x_n, it is easy to attain an abscissa vector $\vec{x} = \{x_m, x_{m+1}, \cdots, x_n\}$, where each element is the x-coordinate of one point on the curve. Substitute elements of $\vec{x}$ in eq. (6) one by one, then the y-coordinate of each point will be gained. In this way, the curve can be reconstructed. In order to make the connection between two neighboring curves perfect and guarantee the contour closed, a generic algorithm of dilation is applied.

For texture parts, a texture synthesis algorithm is performed in the manner of Efros and Freeman's image quilting in [15]. [3] introduces many implementation details such as the patch size w_b. A bigger w_b means better capturing of texture characteristics in the texture patches and thus more similarity between the original texture and the synthesized one. Also, with the help of the texture analysis technique proposed in section 2.3, the scale of the texture elements have been calculated. It is usually assumed that the patch contains one texture element at least.

4 Experimental Results

We implement the proposed image restoration scheme by C code, and test it on some natural images with massive textures. Fig.3 shows results of some important steps of our framework. For image segmentation, the *JSEG* algorithm has 3 parameters that need to be specified artificially [10]. Here, we specify the parameters as follows: TQUAN = -1, NSCALE = -1, threshcolor = 0.8, which make the performance of homogeneity detection perfect. The original image is divided into two regions; see Fig. 4(b).

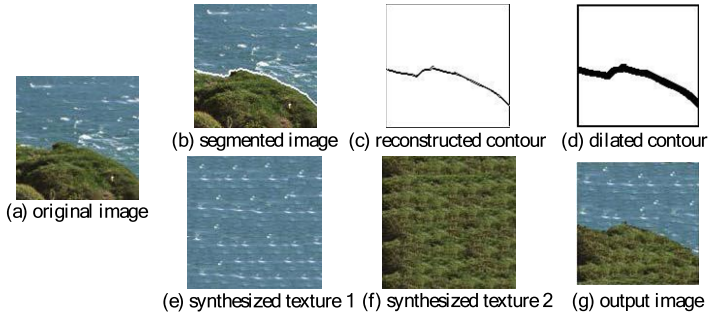


Fig. 3. Some key intermediate results

For contour feature extraction, the split threshold determines how many segments the contour will be partitioned and how many times the iterative fitting should be performed. They assume the inverse relation. In our experiment, the split threshold amounts to 5~10. The degree of polynomial of curve fitting is specified as 4, generally. Fig. 3(c) displays the reconstructed contour. The synthesized texture of upper region in Fig. 3(b) is shown in Fig. 3(e), and the lower one is revealed in Fig. 3(f). Gained the synthesized texture and the reconstructed contour of the region, it is facile to synthesize the final output image displayed in Fig. 3(g).

There are also some final results in Fig.4. Images in the first and third row are original images compressed by JPEG, and the rest are the synthesized ones in our scheme. We can see some distortion between the original image and the synthesized one. That is mainly caused by texture synthesis. However, the distortion of shapes of regions isn't obvious, which proves our shape description method called "piecewise

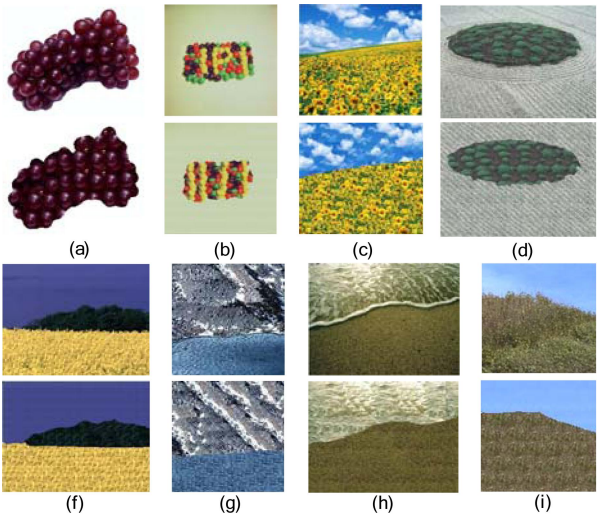


Fig. 4. More image restoration results

iterative curve fitting" effective. If we assume that these massive texture compose the background of one image or video sequences as well as there are some persons, or some horses, *etc*. as the foreground, the distortion in these background regions should not be so visible, as is our motivation of proposing this scheme.

In Table 1, we show the number of bytes used in our proposed framework and compare them to the JPEG compression for the equivalent images. Image encoding is not the goal of our current work. Nevertheless, our synthesized images are fair approximations and we may do more useful attempt for establishing more sophisticated image restoration models to synthesize very realistic images.

Table 1. Comparison of bytes required by JPEG and our scheme (Unit: Bytes)

Image	Fig.4(f)	Fig.4(g)	Fig.4(h)	Fig.4(i)
JPEG	15,872	14,848	28,877	4,506
Our scheme	3,767	4,959	12,666	2,509

5 Conclusion

In this paper, we have investigated a framework based on the human vision characteristic for image restoration by defining a representation/reconstruction scheme for the contours and textures of regions. Piecewise iterative curve fitting method is applied for the contours of texture regions so that contours of regions are split into little curve segments and the shape information can be described by a set of polynomial coefficients. We perform image quilting algorithm for synthesizing inner textures of regions. Our experiments show that the proposed image restoration scheme can get more coding gain but remain vast information people care.

Acknowledgments. This work has been supported in part by the National Science Foundation of China under award No. 60572041.

References

1. Sikora, T.: Trends and Perspectives in Image and Video Coding. Proceedings of the IEEE 93(1), 6–17 (2005)
2. Ndjiki-Nya, P., Makai, B., Blattermann, G., Smolic, A., Schwarz, H., Wiegand, T.: Improved H.264/AVC Coding Using Texture Analysis and Synthesis. In: Proceedings of IEEE International Conference on Image Processing, vol. 2, pp. III-849–852 (2003)
3. Liang, L., Liu, C., Xu, Y.: Real-Time Texture Synthesis by Patch-Based Sampling. ACM Transaction on Graphics 20(3), 127–150 (2001)
4. Efros, A.A., Leung, T.K.: Texture Synthesis by Non-Parametric Sampling. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 2, pp. 1033–1038 (1999)
5. Datsenko, D., Elad, M.: Example-Based Single Document Image Super-Resolution: A Global Map Approach with Outlier Rejection. J. Multidimen. Syst. Signal Processing 2–3, 103–121 (2007)
6. Elad, M., Datsenko, D.: Example-Based Regularization Deployed to Super-Resolution Reconstruction of A Single Image. Comput. J. 18(2–3), 103–121 (2007)
7. Multi-Resolution Recursive Peak Tracking: An Algorithm for Shape Representation and Matching. Dept. of Electrical and Computer Engineering, Carnegie-Mellon University, `http://www.ece.cmu.edu/research/publications/1984/CMU-ECE-1984-029.pdf`
8. McClure, D.E., Vitale, R.A.: Polygonal Approximation of Plane Convex Bodies. J. Math. Anal. Appl. 51, 326–358 (1975)
9. Albano, A.: Representation of Digitized Contours In Terms of Conic Arcs and Straight Line Segments. Computer Graphics and Image Processing 3, 23–33 (1974)
10. Deng, Y., Manjunath, B.S.: Unsupervised Segmentation of Color-Texture Regions in Images and Video. IEEE Trans. on Pattern Analysis and Machine Intelligence 23(8), 800–810 (2001)
11. Jing, F., Li, M., Zhang, H.: An Efficient and Effective Region-Based Image Retrieval Framework. IEEE Transactions on Image Processing 13(5), 699–709 (2004)
12. Adamek, T., O'Connor, N.: Efficient Contour-Based Shape Representation and Matching. In: MIR 2003, pp. 138–143 (2003)
13. Pavlidis, T.: Algorithms for Graphics and Image Processing. Computer Science Press (1982)
14. Chen, C.H., Pau, L.F., Wang, P.S.P.: The Handbook of Pattern Recognition and Computer Vision, 2nd edn. World Scientific Publishing Co., Singapore (1998)
15. Efros, A.A., Freeman, W.T.: Image Quilting for Texture Synthesis and Transfer. ACM SIGGRAPH, 341–346 (2001)

A Real-Time NURBS Interpolator with Feed Rate Adjustment

Liyan Zhang[1,*], Kuisheng Wang[1], Yuchao Bian[2], and Hu Chen[2]

[1] School of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing, China, 100029
zhangly@mail.buct.edu.cn
[2] Beijing Shouke CATCH Electric Technology Co., Ltd., Beijing, China, 102200

Abstract. When realizing a real-time NURBS curve interpolator in CNC machining with high speed and high accuracy, many practical problems are encountered. Algorithms based on Taylor's expansion are developed for realizing NURBS curve interpolator that takes into account the feed rate adjustment to the curvature radius of NURBS curves. By means of the flexible acceleration/deceleration (ACC/DEC) planning, the velocity profile can be chosen dynamically to meet the requirements of the machining process. In addition, a new strategy based on the geometrical properties of NURBS curves is suggested to predict the deceleration point. Therefore, the process of acceleration and deceleration is so smooth that can satisfy the need of machining tool dynamic properties. This NURBS interpolator has been validated on a CNC system, which improves the machining accuracy for both position and velocity control.

Keywords: NURBS, Interpolate, Acceleration/deceleration, CNC.

1 Introduction

Non-Uniform Rational B-Spline(NURBS) curve has been used in CAD systems for a while because it offers an exact uniform representation of both analytical and free-form parametric curves. ISO10303 (STEP) uses NURBS as its foundational geometry representation scheme for free-form curves and surfaces, which facilitates the data exchange between different CAD/CAM systems [1]. On the other hand, conventional CNC systems support only motion along lines and circular paths. Any other tool paths should be approximated with piecewise line segments by CAM systems. The approximation produces errors between the desired path and the commanded path. In order to minimize contour errors, a great number of tiny line segments are used to approximate the curve. However, these tiny segmented contours have drawbacks due to the increasing necessity for faster machining speed and for the machining of complex part shapes.

It's necessary to develop a real-time curve interpolator [2-4]. Using NURBS interpolation results in a reduction of program size of $1/10$ to $1/100$ compared to the size of a comparable linear interpolation part program and significantly improves the fundamental

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1064 – 1071, 2008.
© Springer-Verlag Berlin Heidelberg 2008

accuracy issue [5]. Shopitalni and Koren presented a real-time interpolation algorithm to generate parametric curves for three-axis machining that maintains a constant feed rate along the cut [6]. Moreover, the contour errors caused by this interpolator are much smaller than those caused by the conventional curve approximation. Later, many kinds of parametric curve interpolation algorithms have been proposed. Among these parametric curves, NURBS is currently attracting the most attention. M.Y.Cheng and M.C.Tsai compared different numerical algorithms for implementing the real time NURBS curve interpolators with constant feed rate [7]. Their works indicate that the algorithm based on Taylor's expansion is feasible for implementing the real time NURBS curve interpolator.

In recent years, since the tremendous progress made in the development of tool materials and machine tool design, high speed machining has been widely used. When machining free-form surface with a high speed, because of the curvature radius changing drastically, the surface might be gouged or the cutting tool damaged. CNC systems should compromise between accuracy and efficiency. A lower feed rate was usually assigned by traditional CNC to prevent gouging or tool damage. Therefore, different variable feed rate interpolators have been studied [8-9]. C.W.Cheng and M.C.Tsai reported that acceleration/deceleration before feed rate interpolation was developed to achieve highly accuracy [10]. Nevertheless they did not mention such practical problems as predicting the deceleration period.

In the paper we focus on the practical problems to be solved when implementing NURBS curve interpolator on CNC systems. The conflict of machining speed, machining accuracy and system efficiency is dealt with. A simple but effective method is suggested to adapt the machining speed to the curvature radius of NURBS curves. Furthermore, a new strategy based on the geometrical properties of NURBS curves is suggested to predict the deceleration point. Therefore flexible ACC/DEC planning is realized, which makes the NURBS interpolation much smoother during startup and brake.

2 NURBS Curve Representation

A NURBS curve can be represented parametrically by Equation (1):

$$P(u) = \frac{\sum_{i=0}^{n} \omega_i d_i N_{i,k}(u)}{\sum_{i=0}^{n} \omega_i N_{i,k}(u)} \tag{1}$$

Where $N_{i,k}(u)$ is a blending function defined by recursive formula:

$$N_{i,0}(u)\begin{cases}1 & u \in [u_i, u_{i+1}] \\ 0 & u \notin [u_i, u_{i+1}]\end{cases}$$

and

$$N_{i,k}(u)=\frac{(u-u_i)}{(u_{i+k+1}-u_i)}N_{i,k-1}(u)+\frac{(u_{i+k}-u)}{(u_{i+k}-u_{i+1})}N_{i+1,k-1}(u)$$

Where u is the parameter, and $U=[u_0,u_i,...,u_{n+k+1}]$ is the knot vector. $d_i(i=0,1,...,n)$ is the control point. $\omega_i(i=0,1,....,n)$ is the weight factor.

Therefore, given the control points $d_i(i=0,1,...,n)$, the weight factors $\omega_i(i=0,1,....,n)$, the order k and the knot vector $U=[u_0,u_i,...,u_{n+k+1}]$, a k-order NURBS curve can be defined. In practice, when defining the knot vector $U=[u_0,u_i,...,u_{n+k+1}]$, suppose $u_0=u_1=\cdots=u_k$, $u_{n+1}=u_{n+2}=\cdots=u_{n+k+1}$ and $u_0=0,u_{n+k+1}=1$ [1].

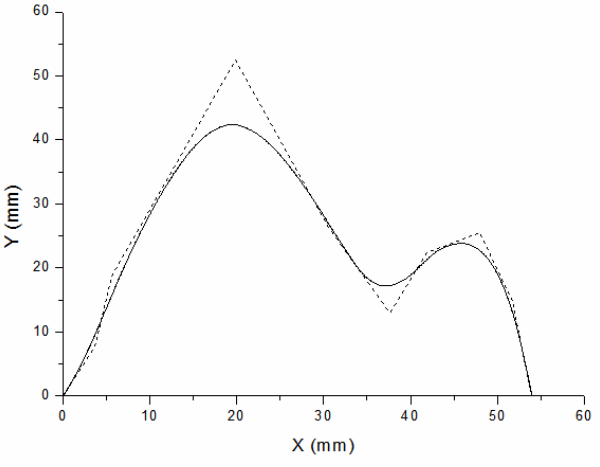


Fig. 1. A NURBS curve

Fig. 1 shows an example of a third-order NURBS curve with eleven control points. The knot vector is （0,0,0,0,0.197,0.564,0.61,0.699,0.765,0.883,1,1,1,1）.

3 Implementation of the Real-Time NURBS Curve Interpolator

3.1 Interpolation on CNC Systems

The interpolation in a CNC system means to calculate the next point continuously. The direct approach for parametric curves interpolator might be increasing fragment of the parameter u , in other words a small uniform increment Δu . Calculate the next point on the curve at each sample time period T. Since the control parameter in this process is u not the time t, the machining feed rate or end-effector's velocity is not

controlled. The key problem of real-time parametric curve interpolation is that the segmentation should be based on curve segment of expected length Δs rather than equal increment of parameter Δu .

According to differential geometry, the desired feed rate v is defined by the following equation:

$$V(t) = \frac{ds}{dt} = \left(\frac{ds}{du}\right)\left(\frac{du}{dt}\right)$$
$$\frac{du}{dt} = \frac{V(t)}{ds/du} = \frac{V(t)}{\sqrt{(\dot{x})^2 + (\dot{y})^2 + (\dot{z})^2}} \tag{2}$$

Equation (2) describes the relationship between the geometric properties of NURBS and the motion properties of the machine tool. A computationally efficient solution of equation (2) is necessary so as to realize a real-time interpolation for parametric curves. However, because the solution of equation (2) is difficult to obtain in general case, approximation forms are used based on Taylor's expansion.

If the curve has the small radius of curvature, a second order approximation is usually adequate, hence:

$$u_{i+1} = u_i + T\frac{du}{dt}\Big|_{u=u_i} + \frac{T^2}{2}\frac{d^2u}{dt^2}\Big|_{u=u_i} \tag{3}$$

where T is the sampling time. If the curve has a large radius of curvature, a first order approximation is used.

$\frac{d^2u}{dt^2}$ and $\frac{du}{dt}$ can be obtained by de Boor recursion algorithms[4].

Equation (3) prescribes how the value of u_{i+1} can be calculated on the basis of the current value of u_i . If the new value of u_{i+1} can be obtained, the next point $(x_{i+1}, y_{i+1}, z_{i+1})$ along NURBS curve can be calculated [4].

3.2 Contour Errors Analysis and Feed Rate Adjustment

Although each interpolation points are located at the NURBS curve, there are also contour errors h . The contour error can be calculated from below equation:

$$h = \frac{(FT)^2}{8r} \tag{4}$$

where F is the machining feed rate. T is the system sampling time and r is the radius of curvature at that point.

In order to analyze the contour errors caused by the real-time NURBS interpolator, the example shown in fig. 1 is used as an illustration. The second order approximation is used. The simulation results of the contour errors are plotted in fig. 2 where the machining feed rate is 6m/min and the sampling time T is 1ms. The maximum contour error in

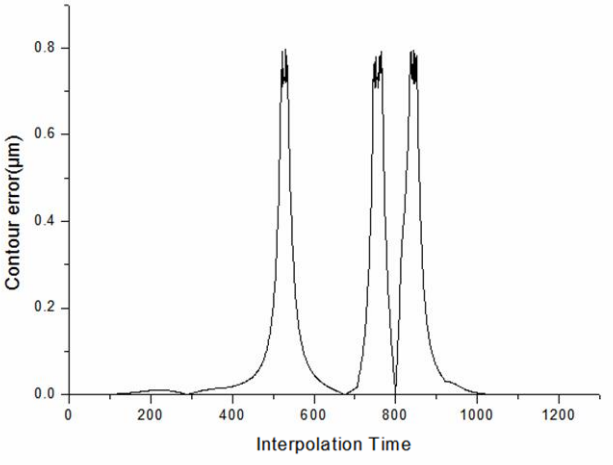


Fig. 2. The results of the contour errors

this example is 0.798 μm , which is located at the point where the curve has the minimum radius of curvature. Another interesting result can be obtained from fig. 2. There are three inflexions in the curve shown in fig. 1. Therefore, there are three peaks in fig. 2 that are caused by small radius of curvature.

Usually, the sampling time is fixed in a system. If the feed rate becomes higher and remains constant, the maximum error might not meet the machining accuracy requirement. The principle is "The machining accuracy is first", when dealing with the conflict of machining accuracy and speed. Therefore, the contour error should be monitored in real time; and if it is larger than the prescribed tolerance, the feed rate will decrease to adjust the radius of curvature.

Hong-tzong, Yau and Ming-jen Kuo optimized the feed rate based on the curvature of NURBS curve and machine dynamic response [2]. It might be difficult to calculate complex mathematical operations in real time. Hence, a practical method, which is directly changing the original increment of parameter Δu according to the geometric properties of NURBS curves and the rule, can be used to change the interpolation distance in order to achieve the feed rate adjustment.

3.3 Flexible Acceleration/Deceleration (ACC/DEC) Planning

3.3.1 The Principle of Flexible ACC/DEC Planning

At the beginning and the end of machining, CNC machine tools can not move with the velocity from 0 to a desired feed rate or from a given feed rate to 0. On traditional CNC systems, a certain velocity profile must be used to obtain smooth motion. Different velocity profiles (i.e. trapezoidal, exponential and bell shape) are studied to satisfy requirements of the machining process [11]. However, changing the velocity profile should change the control system. Flexible ACC/DEC planning means that the velocity profile can be chosen dramatically since the description of different velocity

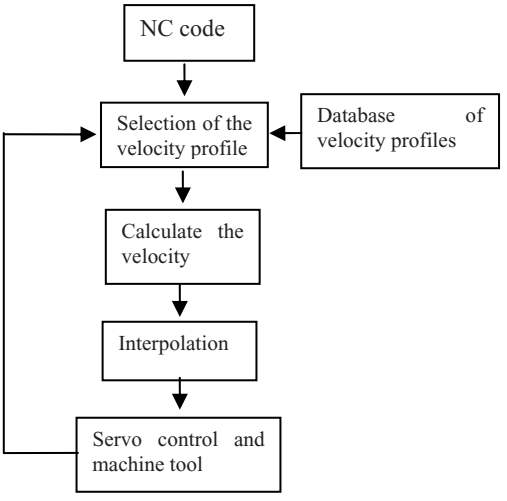


Fig. 3. The procedure of flexible ACC/DEC planning

profiles is saved separately from main control system. The procedure of CNC system with flexible ACC/DEC planning is shown in the fig.3. Various velocity profiles are digitalized, and then saved as tables in a database. Among them, the CNC system selects the suitable velocity profile according to the machining file and the mechanic characteristics of the machine tool.

The procedures for implementing the NURBS curve interpolator with flexible ACC/DEC planning are summarized below:

1. Calculate the velocity of every sampling time T according to the chosen velocity profile, then save them in the ACC velocity table and the DEC velocity table.
2. Judge where the interpolate point is the acceleration period, the constant period or the deceleration period. Get the velocity from the relevant table.
3. Calculate Δu , P_{i+1} and the contour tolerance h. Ensure h $\leqq$ the desired tolerance ε .

In the acceleration period, a CNC system can easily calculate a velocity table according to the given feed rate and the type of velocity profile. As for the deceleration period, the first and important problem is how to judge the start point of the deceleration period, which is the deceleration point.

3.3.2 Judgment of the Acceleration Point and the Deceleration Point

As known from the above discussion, the key problem is how to calculate the acceleration point and the deceleration point especially the deceleration point. Actually, it is required to calculate the parameter u at these points.

The NURBS curve is symmetrical in geometric form with respect to the parameter u , which has been proved by the author [4]. This property is used to produce the deceleration table. In other words, because the same velocity profile is often adopted in acceleration and deceleration period, the data in the deceleration table is exactly the

same as the data in the acceleration table, just in reversed order. This means that the acceleration point of the reversed NURBS curve is the deceleration point of this NURBS curve. Thus, we can obtain the DEC point by means of the ACC point of the reversed curve.

4 Simulating Feed Rate Fluctuation

The NURBS curve shown in fig. 1 is also used to evaluate the performance of feed rate fluctuation. As discussed before, the bell velocity profile is selected to control feed rate, which carries out before feed rate interpolation. The accuracy is also controlled and the desired contour error ε is 0.8 μm . The feed rate is 140 mm/s. From the start and the end of the graph of the fig. 4, the bell velocity profile can be seen. During the middle of this graph, three are three points with feed rate fluctuates. Because there are three inflexions in the curve shown in fig. 1, with smaller radius of curvature, the contour error is larger than the desired error, so the feed rate is automatically reduced to adjust the radius of curvature.

Therefore, under high speed machining with the proposed NURBS interpolation, better machining accuracy and quality are guaranteed although the machining time increases.

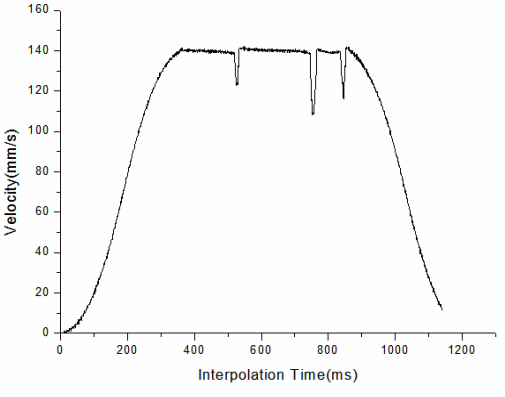


Fig. 4. Feed rate fluctuation

5 Conclusion

In this paper, the NURBS curve interpolator based on the second order Taylor's expansion is implemented with the flexible ACC/DEC planning, where the velocity profile can be chosen dynamically to meet the requirements of the machining process. The DEC point can be obtained by means of the ACC point of the reversed curve. This interpolator also considers controlling the contour error, so the feed rate can adapt to the curvature radius. This NURBS interpolator has been tested in the

CATCH-CNC system and the performance evaluation demonstrated its practical feasibility and its effectiveness.

Acknowledgements. The work reported in this paper was supported by Beijing Municipal Science & Technology Commission under grant no. H030330020110.

References

1. Shi, F.: CAGD & NURBS. Beijing University of Aeronautics & Astronautics Publishing Press, Beijing (1994) (in Chinese)
2. Yau, H.T., Kuo, M.J.: NURBS Machining and Feed Rate Adjustment for High-speed Cutting of Complex Sculptured Surfaces. Int. J. Prod. Res. 39(1), 21–41 (2001)
3. Elkeran, A., El-Baz, M.A.: NURBS Federate Adaptation for 3-axis CNC Machining, http://www.maintenanceresources.com
4. Bian, Y.: Study on Interpolation Algorithm of Free-form Curves and Surfaces in CNC Systems. MSc Thesis. Beijing University of Chemical Technology (2004) (in Chinese)
5. Cheng, C.W.: Design and Implementation of Real-time NURBS Curve and Surface Interpolators for Motion Controllers. PhD Thesis. National Cheng Kung University (2003)
6. Shpitalni, M., Koren, Y., Lo, C.C.: Realtime curve interpolators. Computer-Aided Design 26(11), 832–838 (1994)
7. Cheng, M.Y., Tsai, M.C., Kuo, J.C.: Real-time NURBS Command Generators for CNC Servo Controllers. International Journal of Machine Tools & Manufacture 42, 801–813 (2002)
8. Jung, H.B., Kim, K.: A New Parameterizations Method for NURBS Surface Interpolation. Int. J. Advanced Manufacturing Technology. 16, 784–790 (2000)
9. Farouki, R.F., Tsai, Y.F.: Exact Taylor Series Coefficient For Variable Feed Rate CNC Curve Interpolator. Computer-Aided Design 33, 155–165 (2001)
10. Cheng, C.W., Tsai, M.C.: Real-time Variable Feed Rate NURBS Curve Interpolator for CNC Machining. Int. J. Advanced Manufacturing Technology 23, 865–873 (2004)
11. Zhang, L.: Control of Acceleration and Deceleration in Data Sampling Interpolation. Journal of Beijing University of Chemical Technology 29(3), 91–93 (2002) (in Chinese)

A Genetic Algorithm for Automatic Packing in Rapid Prototyping Processes

Weidong Yang, Weiling Liu, Libing Liu, and Anping Xu

School of Mechanical Engineering, Hebei University of Technology, Tianjin, China
yangweidong@hebut.edu.cn

Abstract. Rapid prototyping (RP) is an emerging technology for creating physical prototypes from three-dimensional computer aided design models by using the additive process with layers. The loading multiple parts into the building space of RP machines, which is a packing optimization problem, is considered as the most important factor to maximize the utilization of space in the building cylinder and reduce building time. To obtain optimized packing result, a genetic algorithm (GA) is developed and implemented. In this paper, chromosomes consist of two segments: one represent the sequence of parts packing by transposition expression, the other represent the rotation mode of parts. This paper will detail the development of a specific algorithm to solve the conversion between the scanning sequence and part layout in two-dimension packing and three-dimension packing of RP processes. Experiments show that the building space is to be used more productively by this packing method for multiple parts.

Keywords: Rapid Prototyping (RP), Genetic Algorithm (GA), Packing.

1 Introduction

Since the life cycle of products is getting shorter due to the rapid industrial development and customers diverse needs, the reduction of the time. Rapid prototyping (RP) technology used from late 1980s has taken its place in CAD/CAM and has been expected to deal with dynamic manufacturing conditions. RP is a material additive process, where a 3D computer model is sliced and reassembled in a real space, layer by layer. Compared with the existing material removal process, RP has the capability of producing prototypes rapidly forming complex shapes, performing the modification process easily and fabricating multi-products.

The building cylinder of different types of RP machines can facilitate more than one part built at a time. This has led to the need of an application to pack this build cylinder. Therefore, the determination of packing is considered as the most important process condition in the RP operation. In the real workshops, operators usually decide this parameter, so the performance and build time of parts mainly depend on their experience level.

2 Packing and GA

To solve the packing problem, Zhu [1] applied two-dimensional manually interactive layout system based on the bounding box approach to the two-dimensional (2D)

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1072–1077, 2008.
© Springer-Verlag Berlin Heidelberg 2008

packing problem. Users input the STL models of three-dimensional (3D) parts needed to be packing, any of the bounding of STL models can be dragged with the mouse to the places arbitrary in the interface. Double-click on a bounding box could make it to be rotating 90 degrees. The packing system can determine their new position by coordinate transformation according to the position change of any of STL models. An example accomplished by the interactive packing system is shown in Fig.1.

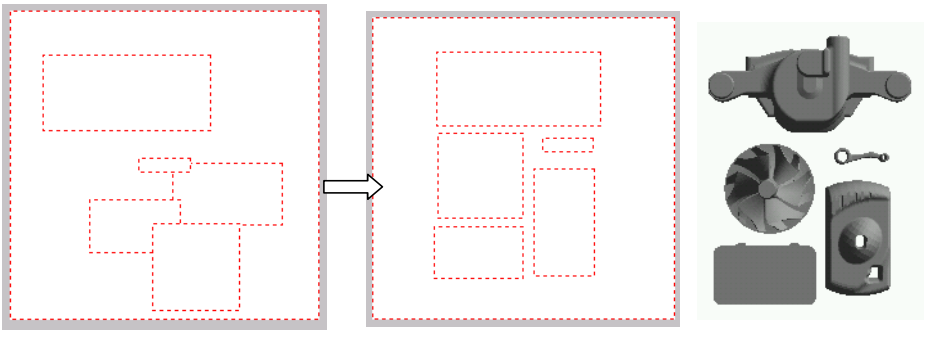


Fig. 1. Example of manually interactive packing

When the number of parts is not too many, the interactive packing system is feasible. However, as the number and complexity of prototypes increase, operators have typically a limited capability in parts packing, including a time consuming and the missing the collision check. The reasonable way should be packed the parts automatically, and then users modify the few inappropriate layout in the interactive interface, the final result is obtained.

The classic bin-packing problem has been shown to be NP-complete problem. Bin-packing has been defined in several different forms, depending on the application [2]. Bin-packing has many aliases, such as stocking cutting, vehicle loading, scheduling and knapsack problem.

Genetic algorithm (GA) is just one of the many exact and heuristic methods used to solve the bin-packing class of problems. To solve the packing problem, Jakobs [3] and Albano and Sapuppo[4] applied GA and BL approach, and the concept of no-fit-polygon, respectively, to the 2D packing problem. Na et al.[5] attempted to simplify and solve the problems associated with 2D shape by transforming them into the quardtree structure. All these researches on the packing problem usually focused on 2D packing problems.

The problem with the 2D representations of the packing problem is that there are many assumptions and constraints need to limit the problem to two dimensions. To solve the 3D packing problem, Szykman and Cagan [6] introduced a simulated annealing (SA) technique as an optimal for a 3D packing problem.

Assumptions about the shape of the parts and constraints as to placement do not exist in three dimensions. Corcoran and Wainwright have previously researched 3D packing problems, as have House and Dagli [7, 8].Their GA application was used to

solve the packing of 3D rectangular parts. A GA application has also been developed for the selective laser sintering (SLS) machines packing problem [9].

The automatic packing can be abstracted to be the problem of layout distribution or combination optimization, which can be solved by the mathematical methodology such as SA, GA and so on. When the bounding boxes of parts are determined, their sizes will be rather different. Therefore, the method of the automatic packing abstracted to be layout distribution subject to certain restrictions, so the solution of the automatic packing is solved by the combination optimization approach in this article.

The automatic packing is abstracted to be the problem of combination optimization. The basic processes of GA for the automatic packing and path optimization such as crossover, mutation, and so on are almost the same, which can refer to [10]. The differences between the packing and path planning are chromosome expression and fitness function. The paper will detail chromosomal decoded method of 2D packing and 3D packing.

3 Decoded Algorithm of Chromosome Expression

3.1 Coded Representation of 2D Packing

The 2D packing for parts is carried in X-Y coordinate. All parts must be placed in the X-Y plane with Z =0. The chromosomes in 2D packing divide into two segments: one represents the packing sequence of parts by transposition expression; the other represents the rotation means of parts (0 – non-rotating; 1 – rotating 90 degree).

Transposition expression can naturally express the scanning sequence of the paths. The scanning sequence of paths must be decoded to the 2D layout of parts, which is accomplished by a specific algorithm. In order to place parts more compact, and decode the chromosome easily, first of all, the maximum length of X-axis (X_{max}) is fixed. This value may be the maximum building dimension in the direction of X-axis of RP machine or may be the one defined by users which is less than the maximum building dimension, which depends on the number of parts, the size of parts, and different RP techniques. The fitness function of chromosome is the maximum length in the direction of Y axis for packing results.

The basic principle of decoded algorithm of 2D packing is to layout parts according to their length in the direction of Y-axis. The length in the direction of Y-axis for the part is small, whose sequence number is ensured to be also small. The procedure of this algorithm is as follows:

1 Determine the numbering of the current part (N_i) having been packing, and then obtain its sizes (X_i, Y_i) of the bounding box.
2 Check whether the part is rotated. If it does, exchange the sizes (X_i, Y_i) of the bounding box.
3 If $X_i > X_{max}$, not packing, interruption.
4 Parts having been packing are sorted by the Y coordinate of the bounding boxes from small to big. By adding 0, ascending array $\{y_0, y_1, y_2, \text{---}, y_{i-1}\}$ is obtained.
5 set j=0.

6 Suppose the minimum length (Y_{min}) in the direction of Y axis is equal to y_j, attempt to pack the part along this length.
7 If may be laid down, this part is arranged in smallest position of X coordinate; Otherwise, $j = j+1$, returns to the 6th step.
8 If the layout of all parts is completed, the end; otherwise $i = i+1$, to the first step.

Because parts must stay between a certain spaces, the bounding box size of parts is larger than their actual size. Fig.2 shows the results of 18 parts using the algorithm for their packing, in which X_{max} is equal to 360 mm, the number of evolution is 100. After packing, the total length for the Y direction is 355 mm, and the area utilization is 95%.

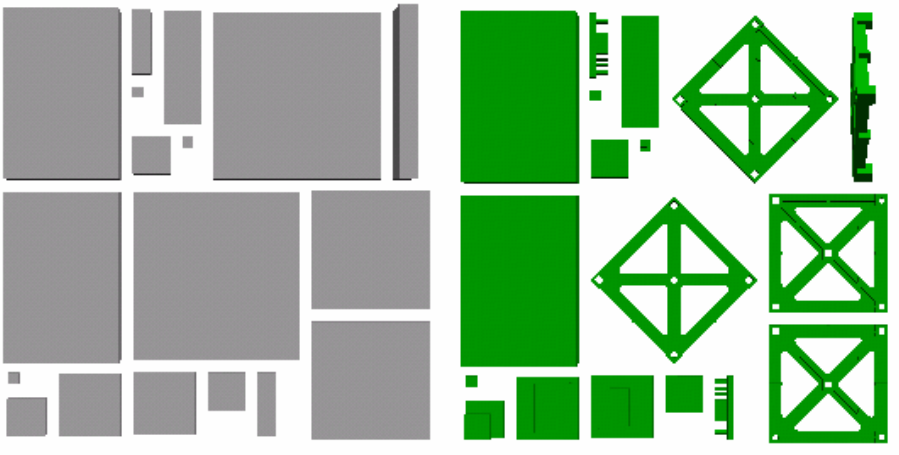


Fig. 2. Example of 2D layout

3.2 Coded Representation of 3D Packing

The 3D packing for parts is layout in XYZ coordinate. All of parts must be placed in the building space of the cylinder. The chromosomes in 3D packing consist of two segments: one represents the parking sequence of parts by transposition expression, the other represents the rotation mode of parts by using random integer from 1 to 6 which correspond the bounding box dimension of the part for XYZ, YXZ, XZY, ZXY, YZX, ZYX respectively.

The transposition representation is decoded for the layout of 3D parts; the decoded algorithm in 3D packing is more complex than that of 2D packing. To display parts as compact, and easy to decode chromosome, firstly the maximum sizes (X_{max} and Y_{max}) in the XY plane are fixed. These values may be the maximum forming sizes of the forming machine in the XY plane, or may be the ones set by users which are less than XY forming sizes, depending on the number of parts, their sizes and techniques. The fitness function of chromosome is the maximum length in the direction of Z-axis for packing results.

The basic principle of decoded algorithm of 3D packing is to layout the parts according to their length in the direction of Z-axis. The length in the direction of Z-axis for the part is small, whose sequence number is ensured to be also small. The procedure of this algorithm is as follows:

1. Determine the number of the current part (N_i) having been packing, and then obtain its dimensions (X_i, Y_i, Z_i) of the surrounding box.
2. Check whether the part is rotated. if it does, exchange the dimensions (X_i, Y_i, Z_i) of the bounding box.
3. If $X_i > X_{max}$ or $X_i > Y_{max}$, not packing, interruption.
4. Parts having been packing are sorted by the Z coordinates of the bounding box from small to big. By adding 0, ascending array { z_0, z_1, z_2, ---,z_{i-1}} is obtained.
5. Set j=0.
6. Suppose the minimum value (Z_{min}) in the direction of Z axis is equal to z_j, attempt to carry on the part for 2D layout to in this Z_{min}.
7. If may be laid down, this part is arranged in this position; Otherwise, $j = j$+1, return to Step 6.
8. If the layout of all prototypes are completed, the end; otherwise $i = i$ +1, to the first step.

Fig.3. illustrates the results of 18 parts(shown in Fig.2) using the algorithm for their 3D packing, in which X_{max} is equal to 200mm, Y_{max} is equal to 250mm, the number of evolution is 100. After packing, the length for the Z direction is 135mm, and the volume utilization is 86%.

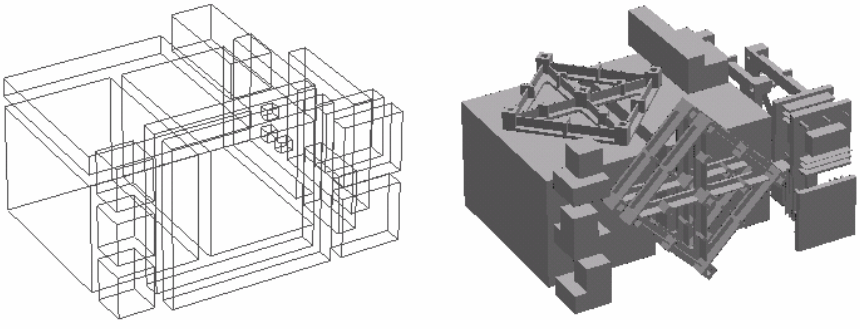


Fig. 3. Example of 3D packing

4 Conclusions

It is meaningless to compare the packing results obtained from automatic approach with those of manual approach, because the efficiency depends on numbers, complexity of parts and operator's experience level. With the developed software the appropriate packing results are obtained within a suitable time under a certain number

and complexity of parts. In addition, it significantly reduces the time to determine the optimal packing, whereas the number of parts increases, the time required becomes very long.

References

1. Zhu, J.: Study on Adaptive Rapid RP-CAPP System. PhD Dissertation, Tsinghua University, Beijing, China (1997) (in Chinese)
2. Dyckhoff, H.: A Typology of Cutting and Packing Problems. Eur. J. Oper. Res. 44, 145–159 (1990)
3. Jakobs, S.: On Genetic Algorithm for the Packing of Polygons. Eur. J. Oper. Res. 50, 165–181 (1996)
4. Albano, A., Sapuppo, G.: Optimal Allocation of Two-dimensional Irregular Shapes Using Heuristic Search Methods. IEEE Trans.Syst. Man Cybernet. SMC 10(5), 242–248 (1980)
5. Na, S.J., Han, G.C., Kim, D.I., Kim, S.L.: A New Approach for Nesting Problem Using Part Decomposition Technology. In: Proceedings of the 23rd International Conference on Industrial Electronics, Control, and Instrument, vol. 3 (1997)
6. Szykman, S., Cagan, J.: A Simulated Annealing Based Approach to Three-dimensional Component Packing. In: Automation 1994, vol. De-69-2, pp. 125–133. ASME, New York (1994)
7. Corcoran III, A.L., Wainwright, R.L.: A Genetic Algorithm for Packing in Three Dimensions. In: Applied Computing Technological Challenges of the 1990, vol. 2. ACM Press, New York (1992)
8. House, R.L., Dagli, C.H.: Genetic Three Dimensional Packer. Intelligent Engineering through Artificial Neural Networks 3 (1993)
9. Ikonen, I., Biles, W.E.: A Genetic Algorithm for Optimal Object Packing in a Selective Laser Sintering Machine. In: Proceeding of the Seventh International FAIM Conference (1997)
10. Yang, W.D., Tan, R.H., Yan, Y.N., Xu, A.P.: The Application of Genetic Algorithm for Scanning Path Planning in Rapid Prototyping. Journal of Computer-aided Design & Computer Graphics 17, 2179–2183 (2005) (in Chinese)

Research of CT / MRI Tumor Image Registration Based on Least Square Support Vector Machines

Zeqing Yang, Libing Liu, Weiling Liu, and Weidong Yang

School of Mechanical Engineering, Hebei University of Technology,
Tianjin 300130, China
zeqing_yang@yahoo.cn

Abstract. In view of the characteristics of the local nonlinear distortion about medical CT and MRI tumor image registration, the method of Least Square Support Vector Machines (LS-SVM) was used to register the images. The verge point of tumor was marked using vector machine weights in this algorithm, and the eigenvalues of tumor images were obtained. The difference of corresponding feature points between CT and MRI tumor image was eliminated adopting the least square algorithm, which not only can effectively remove geometric deformation of the images, but also can be adaptive correct the errors caused by the positioning of feature points. Finally, we had experimented on the Matlab platform, the results show that the algorithm has higher registration accuracy and can meet medical tumor registration requirements. It also play guidance role in image fusion and tumor targeted therapy, which has important application in clinical medicine.

Keywords: Least Square Support Vector Machines (LS-SVM), CT Image, MRI Image, Image Registration, Transformation Model Estimation.

1 Introduction

CT images and MRI images [1] were used to differentiate human body structure in medical diagnosis and treatment. Computed Tomography (CT) has higher spatial resolution and geometrical characteristics, especially to the bone imaging, which can provide a good reference for the lesion localization, but has lower contrast for soft-tissue. Magnetic Resonance Imaging (MRI) is an imaging technique based on image reconstruction of the signals produced by the resonance of the hydrogen protons which abound in the human body in the strong magnetic field. MRI can clearly reflect the soft tissues, organs, blood vessels and other anatomical structures, which is easy to determine the scope of the lesions but not sensitive to calcification point, lacking the rigid bone tissue as positioning reference. For the same patient, the image registration and fusion between CT images and MRI images can improve the diagnosis precision and reliability, making the lesions position more accurately and intuitively display the structures. The fusion of the two kinds of images will not only describe complementary features but also have the possibility of discovering new and valuable

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1078–1086, 2008.
© Springer-Verlag Berlin Heidelberg 2008

information, which makes up for the defect of the images information owing to the different imaging principium.

Medical image registration [2] is to seek a kind of or a series of spatial transform for a medical image to make it match another piece of medical image. Medical image registration has very important clinical values. In recent years, there are many researches on medical image registration, Such as registration based on geometric moment, correlation coefficient and polynomial transforms. These methods need a large amount of calculation and likely to be influenced by noise. The registration and fusion of CT and MRI images is researched on using the pathological changes of the tumor as specimen based on Least Squares Support Vector Machine (LS-SVM) in this work. It has gotten good results on medical image registration of CT and MRI images which have local nonlinear distortion, not only effectively eliminates geometric distortion of the images, but also adaptively corrects the errors due to the low positioning accuracy of the feature points. The experimental results show that the two images can meet the medical tumor registration requirements and also play guidance role in image fusion and tumor targeted therapy, which has important application in clinical medicine.

2 Image Registration

2.1 Image Registration Process

Every registration method [3] can be summarized into the problem of four factors choices including feature space, geometric transformation, similarity measurement and optimization algorithm. Feature space is information set extracted from images and used in registration. When we want to registrate two images, we need carry on a series of geometric transformations to one image. Similarity measurement is actually an energy function, the choice of the extreme value of similarity measurement is very important because it will determine the optimal transformation. Optimization algorithm is the searching strategy in the process of image registration, the best transformation is chosen from many transformations in order to make the similarity measurement can get the optimal value as soon as possible. Image registration is essentially optimizing the transform parameters. According to feature information extracted from images, optimal transformation is achieved through the optimal search strategy in certain searching space, making the registrated images can meet the definition of similarity measurement after transformation, matching with the reference images to a certain degree. The basic frame of image registration is shown in Fig.1.

2.2 The Mathematical Description of Image Registration

From the above analysis, the image registration can be described as a mapping transformation from one data set to another. The two image sets can achieve an optimal match after transformation, that is the error is minimum. Supposed reference image is R and registration image is F , $R(x, y)$ and $F(x, y)$ represents the grey value of reference image and registration image respectively, and then the image registration is expressed as the following formula: $R(x, y) = F(T(x, y))$, where T represents a space geometric transformation.

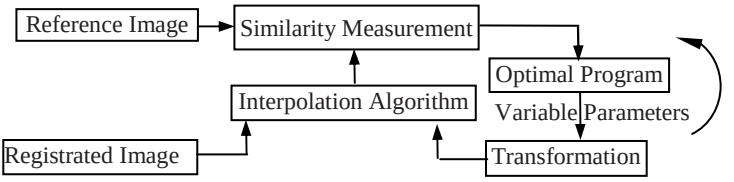


Fig. 1. The basic frame of image registration

2.2.1 Image Feature Extraction Methods

There is no simple point-to-point corresponding relationship between two dislocation images, because the principle of imaging equipment and scanning parameter is different. Firstly, we should extract the image features, such as edge, corner point, line, curvature and so on. The extracted remarkable image features greatly compressed the image information leading to smaller calculation and faster speed. Secondly, the corresponding relationship among the feature point sets is established, and the registration parameters are obtained. The image registration process is generally divided into two stages: one is rough frame stage, which requests to find the feature points of matching objects, not necessarily accurate; the other is convergence stage, which requests to reduce the amount of feature points and find the accurate registration points. The difference between two images is eliminated in the convergence stage using least square matching algorithm.

2.2.2 Image Space Geometric Transformation

The image transformation process from one coordinate system to another is based on a series of or a kind of space geometric transformation [4], which can divide into four kinds, including Rigid Transformation, Affine Transformation, Projective Transformation and Nonlinear Transformation. CT and MRI tumor image registration belongs to multi-mode medical image registration, which can obtain high-resolution bone tissue and soft tissue images. According to the characteristics of CT and MRI tumor image, the Nonlinear Transformation is used to estimate transformation model between two images. The key problem is to determine the mapping function between reference image and registrated image. The mapping function estimation is actually nonlinear function fitting problem.

Supposed the point sets $P = \{p_i\}_{i=1}^{N}$ and $W = \{w_i\}_{i=1}^{N}$ represent N feature points of CT and MRI image respectively. p_i describes a certain feature point (x_i, y_i) in CT image, w_i is a certain feature point (u_i, v_i) in MRI image corresponding to p_i. So p_i and w_i is composed a pair of matching feature point. Supposed point coordinate (x, y) in registrated image corresponds to (u, v) in reference image, it has the following mapping relationship: $\begin{cases} u = f(x, y) \\ v = g(x, y) \end{cases}$.

2.2.3 Similarity Measurement

Similarity measurement [5] [6] is used to describe the similarity between reference image and registrated image. It is an image registration rule reflecting how good or

bad registration of two images. $M(F(T(x,y)),R(x,y))$ is similarity measurement under the transformation T for reference image R and registrated image F. Registration is seeking the transformation to make M can obtain a defined extreme value. The ideal situation of registration is describing the energy function of similarity measurement can obtain maximum or minimum in optimal transformation. It is important that how to select a proper similarity measurement to describe the coincidence degree of geometric position and calculate easily.

2.2.4 Optimization Algorithm

The choice of optimization algorithm is an important step in image registration, because it influences registration accuracy and speed directly. The optimization algorithm LS-SVM is adopted to eliminate difference between two images in the process of image registration.

3 The Image Registration Based on LS-SVM

3.1 The Basic Principle of LS-SVM

The basic idea of Supportting Vector Machine (SVM) [7-11] is that the input data nonlinear map to high-dimensional feature space by mapping function, the regression problems are solved in the high-dimensional feature space [7-8]. Firstly, choose a nonlinear transformation $\phi(\bullet)$, and the sample vector $(x_1,y_1),(x_2,y_2),\cdots,(x_l,y_l)^n$ $x_l \in R^n$, $y_i \in R, i=1,2,\cdots,l$ that is n-dimensional input and l-dimensional output from original space map to the high-dimensional feature space F, the optimal linear regressive function is constructed in this space as following.

$$f(x)=w\bullet\phi(x)+b \tag{1}$$

where $\varphi(\mathbf{x})$ is the mapping function, $\mathbf{w}$ and b can be estimated by the following regularization risk function:

$$\min J(\mathbf{w},b)=\frac{1}{2}\mathbf{w}^T\mathbf{w}+\frac{1}{l}C\sum_{i=1}^{l}L_\varepsilon(y_i,y(\varphi(x_i))) \tag{2}$$

$$L_\varepsilon(y_i,y(\varphi(\mathbf{x}_i)))=\begin{cases}|y_i-y(\varphi(\mathbf{x}_i))|-\varepsilon & |y_i-y(\varphi(\mathbf{x}_i))|\geq\varepsilon\\ 0 & \text{otherwise}\end{cases} \tag{3}$$

Using kernel function of original space to replace the point-product operation in high-dimension space, Equation (4) is gotten.

$$f(\mathbf{x})=\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)K(\mathbf{x}_i,\mathbf{x})+b \tag{4}$$

where $K(\mathbf{x}_i,\mathbf{x}_j)$ ($K(\mathbf{x}_i,\mathbf{x}_j)=\varphi(\mathbf{x}_i)\cdot\varphi(\mathbf{x}_j)$) is the kernel function meeting Mercer theorem, α_i and α_i^* is the Lagrange multipliers.

A new method of support vector machines is presented by Suykens through transforming the the form of VapnikSupport Vector Machine. The loss function is used the least squares linear systems in this method, insteading the quadratic programming method in traditional support vector machines, which has the advantages of computing simple, fast convergence and high accuracy. So it is called Least Squares Support Vector Machines (LS-SVM). Considering the regression estimation problems of SVM, l training datas $(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_l, y_l)$, $\mathbf{x}_i \in R^n$ $y_i = R$ are given, where $\mathbf{x}_i$ is input data, y_i is output data. In the original space, the regression problems can be described as the following form of LS-SVM:

$$\min J(\mathbf{w}, \boldsymbol{\xi}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \gamma\frac{1}{2}\sum_{i=1}^{l}\xi_i^2 \tag{5}$$

Constraint condition is shown as Equation (6) .Where $\mathbf{w}$ is the weight vector, ξ_i is the error variable, b is the deviation value, γ is the adjustable constant.

$$y_i = \mathbf{w}^T\varphi(\mathbf{x}_i) + b + \xi_i, \quad i = 1, 2, 3, \cdots, l. \tag{6}$$

Equality constraints rather than inequality constraints are included in this form. According to equation (5), Lagrangian function is constructed in equation (7). $\alpha_i \in R$ is the lagrangian operator, which can be positive and negative in the form of LS-SVM. According to optimum conditions, the following equation (8) is must met.

$$L(\mathbf{w}, b, \boldsymbol{\xi}, \alpha) = J(\mathbf{w}, b, \boldsymbol{\xi}) - \sum_{i=1}^{l}\alpha_i\left\{\mathbf{w}^T\varphi(\mathbf{x}_i) + b - y_i + \xi_i\right\}, \tag{7}$$

$$\begin{cases} \dfrac{\partial L}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} = \sum_{i=1}^{l}\alpha_i\varphi(\mathbf{x}_i); \dfrac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{l}\alpha_i = 0; \\ \dfrac{\partial L}{\partial \xi_i} = 0 \rightarrow \alpha_i = \gamma\xi_i, i = 1, 2, 3, \cdots, l; \ \dfrac{\partial L}{\partial \alpha_i} = 0 \rightarrow \mathbf{w}^T\varphi(\mathbf{x}_i) + b - y_i + \xi_i = 0 \end{cases} \tag{8}$$

where $i = 1, 2, 3, \cdots, l$, eliminating $\mathbf{w}$ and ξ , the matrix equation (9) is gotten.

$$\begin{bmatrix} 0 & \mathbf{l}_v^T \\ \mathbf{l}_v & \Omega + \dfrac{1}{\gamma}I \end{bmatrix}\begin{bmatrix} \mathbf{b} \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \tag{9}$$

$\mathbf{y} = [y_1, \cdots, y_l]$, $\mathbf{l}_v = [1, \cdots, 1]$, $\boldsymbol{\xi} = [\xi_1, \cdots, \xi_l]$, $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_l]$. According to Mercer theorem, there exists the mapping function φ and kernel function $K(\mathbf{x}_i, \mathbf{x}_j)$, making the equation (10) is gotten.

$$K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i)^T\varphi(\mathbf{x}_j) \tag{10}$$

In this way, the estimated form of LS-SVM is shown as following:

$$y(\mathbf{x}) = \sum_{i=1}^{l}\alpha_i K(\mathbf{x}, \mathbf{x}_i) + b\,. \tag{11}$$

LS-SVM has more extensive practical application [8-10] than the standard SVM, because LS-SVM is simpler and more efficient. Therefore, LS-SVM is used to transformation model estimation in CT/MRI tumor image registration in this paper.

3.2 Transformation Model Estimation with LS-SVM

Supposed f_1 and f_2 are the mapping functions, and meet the expression of $U = f_1(x, y)$ and $V = f_2(x, y)$, where $x = \{x_i\}_{i=1}^N$ and $y = \{y_i\}_{i=1}^N$ are also meet the requirements. The fitness mapping function f_1 and f_2 between the point set P and W can be determined through Minimizing function (5) in order to f_1 and f_2 can pass all feature points in the constraint conditions of function (6). For simplify the problem, f_1 and f_2 is expressed with f .Considering the requirements of CT/MRI tumor images registration in medicine, Gauss RBF Kernel Function is chosen. Its expression is as follows:

$$f(p) = \sum_{i=1}^{N} \alpha_i \exp\left\{-\|p - p_i\|^2 / \sigma^2\right\} + b \tag{12}$$

Where $\|\bullet\|$ indicates Euler's 2-noum of R^2 , that is $\|p - p_i\|^2 = (x - x_i)^2 + (y - y_i)^2$, σ^2 is the kernel parameter, We can get the expressions of α and b through inducing the given feature points into the Equation (9), which is shown as following:

$$\alpha^u = \Omega^{-1}\left(U - \vec{1}b^u\right); \alpha^v = \Omega^{-1}\left(U - \vec{1}b^v\right); b^u = \frac{\vec{1}^T \Omega^{-1} U}{\vec{1}^T \Omega^{-1} \vec{1}}; b^u = \frac{\vec{1}^T \Omega^{-1} V}{\vec{1}^T \Omega^{-1} \vec{1}} \tag{13}$$

Where α^u , α^v , b^u , b^v indicate the parameter α , b of u , v respectively.

$$\Omega = \exp\left\{-\|p_j - p_i\|^2 / \sigma^2\right\} + \gamma^{-1} I \text{ , } U = (u_1, \cdots, u_N), V = (v_1, \cdots, v_N).$$

The mapping function f_1 and f_2 can be gotten after solving the α^u , α^v , b^u , b^v .

3.3 Evaluation Standards of Image Registration

The Mean Error (ME) and Root Mean Square Error (RMSE) are used to evaluate the results of image registration in Experiments. The definition is as follows:

$$ME = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(u_i - tu_i)^2 + (v_i - tv_i)^2} \tag{14}$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \left[(u_i - tu_i)^2 + (v_i - tv_i)^2\right] / n} \tag{15}$$

where (u_i, v_i) is the feature point of the reference image, (tu_i, tv_i) is corresponding feature points mapping from feature point (x_i, y_i) of registrated image to reference image, and n is the number of feature points for training or testing.

4 The Analysis of Experiment

We choose CT and MRI images of the same patient as research objects, and the CT image of the lung is gotten using Siemens SomatomARSP helix CT.The standard parameters are 120kv, 100mA, 512×512; the MRI image is gotten by the magnetic resonance systemic MRI scanner produced by Siemens. The MRI image matrix is 256×256.The two kinds of images are both provided in the form of film, after being processed by HP scanner, and then researched on image registration senting into PC. Fig.2 and Fig.3 is original CT and MRI image respectively, and Fig.2 is taken as reference image and Fig.3 is taken as the registered image.

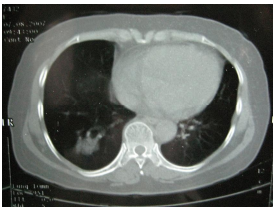


Fig. 2. The original CT image

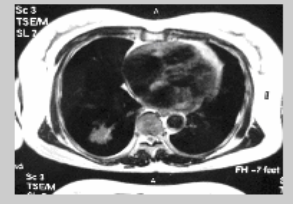


Fig. 3. The original MRI image

Fig. 5. The CT image after choosing feature points

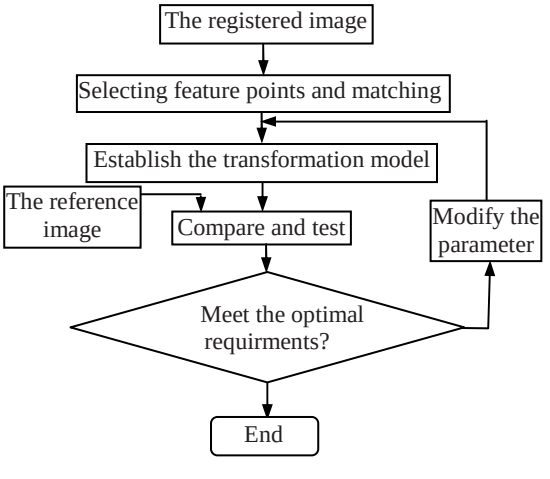


Fig. 4. The flow of image registration

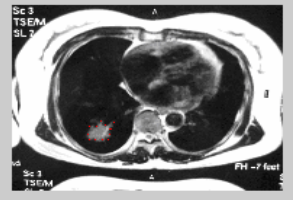


Fig. 6. The MRI image after choosing feature points

Fig. 7. The registeration image

The following experiments are based on LS-SVM flexible registration method, mainly contain four steps,the flow is shown as Fig.4. Firstly, choose a mount pairs of matching feature points from the registered image. 11 pairs of matching feature points are manually chosen by the feature point chosen tool in the Image processing Toolbox provided by Matlab, which is shown as Fig.5 and Fig.6. Secondly, establish the corresponding connection of the pairs of matching feature points of the two images and estimate the transformation model between images based on LS-SVM. And then the whole image is resampled and transformed according to the established transformation model, make the result of the transform carried on the registered image in order to realize the image flexible registration.

Fig.7 is shown as the result of the image registration using the following parameters: $\gamma^u = 3.652\times10^5$, $\sigma_u^2 = 6.356\times10^5$, $\gamma^v = 8.375\times10^8$ and $\sigma_v^2 = 2.345\times10^6$ .Manually choose six pairs of feature points as testing points, using ME and RMSE as appraisal index of the registration to weigh the registration precision, the experimental results are shown as Table 1, which indicates that the target images meet the requirements of the registration.

Table 1. Experiment data

Error	The testing points	The training points
ME	0.3214	0.5625
RMSE	0.4265	0.6742

5 Conclusions

The basic frame structure of image registration methods are analyzed in detail, In view of the characteristics of the local nonlinear distortion about medical CT and MRI tumor image registration, considering SVM has predominance in the field of non-linear pattern recognition of less sample and high-dimension and well versatility, LS-SVM is adopted in the image registration, which not only can effectively remove geometric deformation of the images, but also can be adaptive correctethe errors caused by the positioning of feature points. We had experimented on the Matlab platform, the results show that the algorithm has higher registration accuracy and can meet medical tumor registration requirements. It also play guidance role in image fusion and tumor targeted therapy, which has important application in clinical medicine.

Acknowledgments. This work is supported by the Hebei Province key scientific and technological project (04213505D).

References

1. Wang, M.Q., Shao, B., Ye, J.H.: Application and Prospect of Medical Image Fusion. Information of Medical Equipment 22(5), 48–50 (2007)
2. Maurer, C.R., Aboutanos, G.B., Dawant, B.M.: Registration of 3-D images using weighted geometrical features. IEEE Trans. Med. Imag. 15, 836–849 (1996)

3. Zeng, W.F., Li, S.S., Wang, J.G.: Translation, rotation and scaling changes in image registration based affine transformation model. Infrared and Laser Engineering 30(1), 18–20 (2001)
4. Leventon, M.E., Grimson, E.W.L.: Multi-Modal Volume Registration Using Joint Intensity Distribution. Image Vision Compute 2, 97–110 (2001)
5. Kybic, J.: Fastp parametric elastic Image Registration. IEEE Transactions on Image Processing 11, 1427–1442 (2003)
6. Zitova, B., Flusser, J.: Image registration methods: A survey. Image and Vision Computing 21, 977–1000 (2003)
7. Shu, H.Z., Luo, L.M., Yu, W.X.: A new fast method for computing legendre moments. Pattern Recognition 33, 341–348 (2000)
8. Cao, L.J., Tay Francis, E.H.: Support vector machine with adaptive parameters in financial time series forecasting. IEEE Transactions on neural networks 14(6), 1506–1518 (2003)
9. Vapnik, V.N.: Statistical learning theory. Wiley, New York (1998)
10. Suyken, J.A.K., Vandewalle, J.: Least Squares Support Vector Machine Classifiers. Neural Processing Letters 9(3), 293–300 (1999)
11. Chua, K.S.: Efficient computations for large least square support vecor machine classifiers. Pattern Recognition Letter. 24(1), 75–80 (2003)

A Novel Format-Compliant Video Encryption Scheme for H.264/AVC Stream Based on Residual Block Scrambling

Shu-bo Liu[1,2], Zheng-quan Xu[1,*], Wei Li[3], Jin Liu[1], and Zhi-yong Yuan[2]

[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,Wuhan University, Wuhan, China, 430079
[2] Computer School,Wuhan University, Wuhan, China, 430079
[3] Alcatel-Lucent China Beijing R&D Center, beijing, 100102
xuzq@whu.edu.cn, liu.shubo@163.com

Abstract. For secure stream media applications, block spatial scrambling can be used as a practice video encryption approach. The scrambling approaches commonly used for encryption of H.263/MPEG-4 stream can not be used for H.264/AVC's directly because the neighboring MacroBlocks, residual blocks and codewords of H.264/AVC stream data are context-sensitive. A novel format-compliant residual block spatial scrambling algorithm for H.264/AVC stream encryption is proposed in this article. In the algorithm, the residual blocks are categorized into different groups and each residual block group uses different random shuffling table, which is generated by the chaotic cryptosystem respectively. The experimental results show that H.264/AVC video stream could be encrypted in real-time while the format is kept compliant.

Keywords: H.264/AVC, format-compliant, residual block scrambling, chaotic cryptosystem.

1 Introduction

It is an important branch for spatial permutation to encrypt image pixels directly. L.Qiao proposed a pure permutation algorithm for the first in video encryption [1]. The algorithm uses a random scrambling sequence to permute video bitstream directly, and can endure the attack for cryptanalysis based on frequency characteristics, but can not endure the attack for cryptanalysis based on known plaintext. In addition, due to permute the video bitstream directly, the algorithm destroys video stream structure that makes the encrypted stream does not keep format compliance. In Ref. [3], J.Wen proposed a format compliant video stream scrambling method on spatial codeword. The main idea of the method is to keep the stream format information unchanged, such as starting-up word, marking code, and permuting the compressed stream spatially on basic data unit. The basic scrambling units could be a single fixed-length or variable length codeword, as well as 8 × 8 coefficient blocks or codewords corresponding a block /macroblock(MB). Whereas sometimes a video data package may only contain a

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1087–1094, 2008.
© Springer-Verlag Berlin Heidelberg 2008

few impactful MBs, which always occur in high-bit-rate video communication, blocks or MBs as basic scrambling unit will lead scrambling space too narrow. In this case, computational complexity of the attack analysis is too low to decrypt the scrambling table. So the method is more suitable for codewords as the scrambling basic unit.

Other authors, eg. Hong, GM. Ahn, J. and Socek, D. etc., proposed several encryption methods for H.264/AVC [2][4-6][8][9], but all those algorithms are not format-compliant.

Unlike the past video encoding recommendation such as H.263/MPEG-4, the neighboring MBs, residual blocks and codewords of H.264/AVC stream data are context-sensitive. It means that, in the process of encode and decode of residual blocks, the number of non-zero coefficient and the number of trailing coefficient should be chosen adaptively, as well as the adaptively update of suffix length of non-zero coefficient codeword, depending on its homologous neighboring data. If the location of residual blocks is scrambled randomly, it would lead to the H.264/AVC bitstream format incompliant.

Ahn proposed a scrambling method based on intra block coding in Ref. [5]. Although the authors considered the residual values, their algorithm is still not format-compliant.

In this paper, a new format-compliant method for encrypting H.264/AVC stream based on residual block scrambling is presented. In Section 2, the approach to scramble residual blocks of H.264/AVC is developed. The experiment results and cryptanalysis of the approach are described in Section 3. Finally, conclusions are drawn in Section 4.

2 An Adapted Residual Block Scrambling Algorithm for H.264/AVC

2.1 H.264/AVC Residual Values Coding

In H.264/AVC, the total number of non-zero transform coefficient levels and the number of trailing one transform coefficient levels are derived by parsing for a given codeword. VLC selection of the codeword is dependent upon a variable nC. The current residual values block nC is derived by nA and nB. Except chroma DC coefficient, other residual values block's nC is derived by the left 4×4 non-zero transform coefficient block and the up 4×4 non-zero transform coefficient block of the current MB corresponding value nA and nB (See Fig. 1).

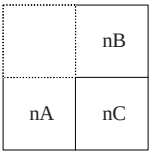


Fig. 1. The current residual values block nC is derived by nA and nB

If scrambling the residual values block directly, two issues would arise: Firstly, an incorrect variable nC would be derived by the changed nA or nB, that would cause the incorrect value of the number of non-zero transform coefficient and the number of

trailing one transform coefficient as lookup the wrong encoding table. This leads to the format incompliant; Secondly, the number of non-zero transform coefficient derived from the permuted residual block would conflict with the number corresponding macroblock_code_mode, luma or chroma coded_block_pattern. This also leads to the format incompliant.

In a word, the scrambling approaches which commonly used for encryption of H.263/MPEG-4 stream can not be used for H.264/AVC's directly that would lead to format incompatibility because the neighboring MBs, residual blocks and codewords of H.264/AVC stream data are context-sensitive.

2.2 The Basic Principle of Algorithm

Based on the analysis of the H.264/AVC video codec characteristics, it is noted that the adjacent block decoder is affected by the value nC of residual block and the number of non-zero coefficient significantly. An adaptive encryption algorithm is proposed. In the algorithm, the residual blocks are categorized into different groups according to non-zero coefficient total_coeff and nC value. Each group uses different scrambling table and finally all the groups are scrambled together.

The basic idea of this algorithm is as below: For each coding slice, the residual blocks with the same non-zero coefficient total_coeff and nC value are divided into the same group. According to the difference of the total_coeff and nC value classified into different scrambling groups. The numbers of scrambling groups is corresponding to the combinations of total_coeff and nC value. Finally, the scrambling groups will be permuted.

The range of non-zero coefficient total_coeff of 4×4 residual block is [0, 16], there being 17 kinds of values. The range of nC value is [-1, 16], there being 18 kinds of values. Therefore, if all residual blocks with different attributes need to be scrambled, there is the requirement of setting the 17×18 scrambling groups. In different applications, to go along with confidentiality requirements, scrambling groups may be appropriately reduced so as to reduce the complexity of the procedure. For example, nC value, in most cases, is -1, 0, 1… 8 and total_coeff in most cases is within the scope of [0, 9]. It can be only take [-1, 8] for nC within 10 values and [0, 9] for total_coeff within 10 kinds of combinations.

The classification of scrambling groups can guarantee that the scrambling must be done only in the group, thus further ensure format compliance of the encoded bitstream.

The scrambling table of each group is generated respectively by a random sequence. For example, it would be obtained through a shuffling algorithm from a pseudo-random sequence generated by the chaotic cryptosystem.

2.3 Algorithm Implementation

Because of residual blocks are categorized into different groups, and each group is scrambled independently through different scrambling table, the processing group classifications and block scramble in a slice would be carried out repeatedly. That results in greatly increasing of the complexity of the calculation. In order to reduce the complexity of the calculation, each residual block is numbered altogether according to the order of its initial address in a slice. Then generate a scrambling table from the

random sequence generated by the chaotic cryptosystem for each scrambling group. The size of the table is related to the number of residual block contained. The scrambling tables of each group will be integrated into a general scrambling table. Afterwards, all residual blocks can be scrambled in a slice through the general scrambling table. It not only ensures that all residual blocks permutated are within the scrambling group they belonged to, that is necessary for the algorithm to keep format compliant, but also ensures all residual blocks are permutated with just one scrambling processing, that offers the high efficiency of the algorithm.

For example, in Fig.2, a section code stream in a slice is divided into A, B, C, and D four Scrambling groups. There are three residual coefficient blocks {A1, A2, A3} in group A, generating scrambling table {A2, A1, A3}. There are three residual coefficient blocks {B1, B2, B3} in group B, generating scrambling table {B3, B2, B1}. There are two residual coefficient blocks {C1, C2} in group C, generating scrambling table {C2, C1}. There are two residual coefficient blocks {D1, D2} in group D, generating scrambling table {D2, D1}. Each block is labeled with a number according to its order in the stream and an initial numbered list is generated. The initial numbered list and its corresponding block list are shown the first row and the second row in Fig.2. Then permute order number of the block member of the group according to the scrambling table of each group respectively, to attain the general scrambling table which is shown in the third row in Fig.2. Finally, permute the blocks in stream through the general scrambling table as the output of ciphertext which is shown in the fourth row.

Original sequence	1	2	3		4	5	6	7	8	...	9	10	...
residual blocks	A1	B1	A2		C1	D1	B2	D2	B3	...	A3	C2	...
scrambling matrix	3	8	1		10	7	6	5	2	...	9	4	...
scrambling residual blocks	A2	B3	A1		C2	D2	B2	D1	B3	...	A3	C1	...

Fig. 2. An example of H.264/AVC residual blocks

The method of scrambling residual blocks can perturb the video image in vision and keep it format compliant, it can be used independently and also can be used a complementary role integrated with other selective encryption algorithm to improve the confidentiality.

3 Experiment Results

3.1 Experiment Environment

(1) PC: Pentium IV-3.2GHz with 512 MB RAM.
(2) Files: Foreman and mobile video files with cif and qcif size.
(3) Algorithm: Scrambling matrix using chaotic sequence. The proposed algorithm's scrambling matrix is independence, and can be chosen freely. In the experiment,

based on the Logistic map $\chi_{n+1}=\mu\chi_n(1-\chi_n)$, a lookup table is used. The results of encrypted bitstream show that it's not only format compliant and secure, but also in real-time.

3.2 Visual Effects

The proposed scheme (II) only scrambling the residual blocks and the scheme (III) of the combination of the scrambling residual blocks and selective encryption are adopted on the foreman and mobile sequences. The two experiment results of two sequences shown in Fig.3.

From Fig. 3, it can be seen that only adopting residual block Scrambling the image becomes blurred, color confusion, but as the increase of I frame intervals, the hided effect of images have been enhanced and we can also distinguish images outline. The encryption of scrambling residual block combined with selective encryption can be very good visual effect [7][10] of secrecy, whether I frame or P frame are difficult to distinguish any content information.

(a) I frame (b) 37th frame (c) I frame (d) 37th frame

(I) Foreman& mobile original frames

(e) I frame (f) 37th frame (g) I frame (h) 37th frame

(II) Scrambling residual blocks

(i) I frame (j) 37th frame (k) I frame (l) 37th frame

(III) The combination of the scrambling residual blocks and selective encryption

Fig. 3. The visual effects of foreman and mobile video files using different encryption

3.3 Cryptanalysis

In Fig.3, it appears that only scrambling residual blocks have not high confidentiality so as to not be used alone in need of high confidentiality occasion. However, critical information in bitstream as a random sequence is used to generate a chaotic table in the process of residual blocks scrambling encryption. The random sequence is similar to a true random sequence to ensure the security of the algorithm. In addition, if the attack of exhaustive analysis applies to the residual scrambling blocks. Suppose that the range of nC is [-1, 8] and the range of total_coeff is [0, 9], which are drawn as a scrambling scope. There are 100 such scrambling groups. The number of scrambling residual blocks is 304 on average in QCIF image of foreman through statistics. In this case, the average foreman Scrambling group has about three residual blocks. Distribution of residual blocks in each scrambling group is in the entire slice. The function of a limited number of Scrambling Groups within only exhaustive analysis is minimal. Additional, through permutations and combinations decrypting certain a scrambling group needs to put residual blocks into image to judge whether it is correct through analysis. And if the blocks around the residual block have not been decrypted, it is obviously unable to determine which permutation and combination is in accordance with the original plaintext. Therefore, the relationship of scrambling groups is not a simple linear. If only 1/3 residual blocks of foreman ciphertext sequence are analyzed through exhaustive analysis, the computational complexity can reach to $(p_3^3)^{33} \geq 2^{84}$. If all residual blocks of the sequence are analyzed through exhaustive analysis, the computational complexity can reach to $(p_3^3)^{100} \geq 2^{252}$, so it is unable to decrypt the image through exhaustive analysis.

Because attack analysis of video images and other information need people to participate and judge, in the circumstances of only ciphertext attack, the number of attack time will be increased in quantities. The security of this algorithm will be further enhanced.

If this algorithm combined with other encryption algorithm is used, their confidentiality will greatly be enhanced [7][10].

3.4 Format Compliant

In principle, this algorithm has maintained unchangeable on syntactic format information, it does not destroy contextual related information relied on residual blocks in H.264/AVC based on context in the process of adaptive encoder, and it is all compatible with H.264/AVC.

In the experiment, the sequences, such as, the foreman and mobile of the encrypted bitstream can be fluently decoded, but only the content of the encrypted image can not be identified.

3.5 Computational Complexity

Computational complexity of residual blocks scrambling mainly focuses on the sequence cipher algorithm to generate the scrambling code table.

Because of the limited number of residual blocks in encoder slice, each slice only need to acquire a small number of random bits from random sequence generator to generate scrambling code table. Therefore, the computational complexity of the method of scrambling residual blocks is small.

The statistics using two types of encryption methods on CIF and QCIF format of the two same sequences is given, and the proportion of H.264 coding time is such as Table 1. After joining encryption algorithm delay is not more than 3ms on average, so it does not affect the real-time encoding and transmission of H.264/AVC.

Table 1. The time ratio of encryption/encode

File	Only Scrambling Residual Blocks	Scrambling residual blocks and selective encryption scheme are used
Foreman_qcif	0.56%	4.87%
Mobile_qcif	0.67%	6.98%
Foreman_cif	0.60%	5.72%
Mobile_cif	0.73%	7.85%

3.6 Impact on Video Compression Ratio

Firstly, encryption process of this algorithm is independent of the video codec and has no impact on the characteristics of the various encoding and decoding. At the same, it does not affect the video image quality and compression ratio. In addition, since the scrambling blocks just only are shifted in location order of bitstream in a slice and do not send any other additional information. Therefore, it will not generate additional data and impact on video image compression ration.

4 Conclusion

Directly scrambling MBs, blocks or codewords for encrypting H.264/AVC stream can not keep format compliant. But if residual blocks are divided into different groups according to the number of non-zero coefficient and value of nC, and each residual block group uses different random shuffling table for scrambling respectively, the scrambling algorithm can keep the original contextual correlation of residual blocks. Accordingly, the algorithm can keep format compliant.

The experimental results show that the encryption algorithm of scrambling residual blocks can disturb the video image in real-time for H.264/AVC video stream while keep its format compliant. The algorithm can be used independently, and also can be used a complementary role integrated with other selective encryption algorithm to improve the confidentiality.

Acknowledgments. The work was supported by the Project (No. 2006CB303104) of the National Basic Research Program (973) of China.

References

1. Qiao, L., Nahrstedt, K.: Comparison of MPEG Encryption Algorithms. International Journal on Computers & Graphics 22, 437–448 (1998)
2. Wen, J., Severa, M., et al.: Fast Self-synchronous Content Scrambling by Spatially Shuffling Codewords of Compressed Bitstreams. In: IEEE International Conference on Image Processing, pp. 169–172 (2002)
3. Wen, J., Severa, M., et al.: A Format Compliant Configurable Encryption Framework for Access Control of Video. IEEE Tran. Circuits & Systems for Video Technology 12, 545–557 (2002)
4. Hong, G.M., Yuan, C., Wang, Y., et al.: A Quality-controllable Encryption for H.264/AVC Video Coding. In: Advances in Multimedia Information Processing-PCM 2006, Processing, vol. 4261, pp. 510–517 (2006)
5. Ahn, J., Shim, H.J., Jeon, B., et al.: Digital Video Scrambling Method Using Intra Prediction Mode. In: Advances in Multimedia Information Processing-PCM 2004, Processing, vol. 3333, pp. 386–393 (2004)
6. Socek, D., Culibrk, D., Kalva, H., et al.: Permutation-based Low-complexity Alternate Coding in Multi-view H.264/AVC. In: 2006 IEEE International Conference on Multimedia and EXPO-ICME 2006 Proceedings, vol. 1-5, pp. 2141–2144 (2006)
7. Xu, Z.Q., Yang, Z.Y., Li, W.: An Overview of Encryption Scheme for Digital Video. Geomatics and Information Science of Wuhan University 30, 570–574 (2007)
8. Thomas, W., Sullivan, G.J., et al.: Overview of the H.264/AVC Video Coding Standard. IEEE Transaction on Circuits and System for Video Technology 13 (July 2003)
9. Fan, Y.B., Wang, J.D., et al.: A New Video Encryption Scheme for H.264/AVC. In: Ip, H.H.-S., Au, O.C., Leung, H., Sun, M.-T., Ma, W.-Y., Hu, S.-M. (eds.) PCM 2007. LNCS, vol. 4810, pp. 246–255. Springer, Heidelberg (2007)
10. Li, W., Xu, Z.Q., Yang, Z.Y., Liu, S.B.: Format-compliant H.264/AVC Video Encryption Algorithm. Journal of Huazhong University of Science and Technology(Nature Science Edition) 35, 13–17 (2007)

A Prototype of Multimedia Metadata Management System for Supporting the Integration of Heterogeneous Sources

Tie Hua Zhou, Byeong Mun Heo, Ling Wang, Yang Koo Lee, Duck Jin Chai, and Keun Ho Ryu*

Database/Bioinformatics Laboratory, School of Electrical & Computer Engineering, Chungbuk National University, Chungbuk, Korea
{thzhou,muni4344,smile2867}@dblab.chungbuk.ac.kr
{leeyangkoo,djchai2520, khryu}@dblab.chungbuk.ac.kr

Abstract. With the advances in information technology, the amount of multimedia metadata captured, produced, and stored is increasing rapidly. As a consequence, multimedia content is widely used for many applications in today's world, and hence, a need for organizing multimedia metadata and accessing it from repositories with vast amount of information has been a driving stimulus both commercially and academically. MPEG-7 is expected to provide standardized description schemes for concise and unambiguous content description of data/documents of complex multimedia types. Meanwhile, other metadata or description schemes, such as Dublin Core, XML, TV-Anytime etc., are becoming popular in different application domains. In this paper, we present a new prototype Multimedia Metadata Management System. Our system is good at sharing the integration of multimedia metadata from heterogeneous sources. This system enables the collection, analysis and integration of multimedia metadata semantic description from some different kinds of services. (UCC, IPTV, VOD and Digital TV et al.)

Keywords: Multimedia Metadata Management Systems, Metadata, MPEG-7, TV-Anytime.

1 Introduction

There is an increasing demand toward multimedia technology in recent years. MPEG-7 is an XML-based multimedia meta-data standard, which proposes description elements for the multimedia processing cycle from the capture (e.g., logging descriptors), analysis/filtering (e.g., descriptors of the MDS, Multimedia Description Schemes), to the delivery (e.g., media variation descriptors), and interaction (e.g., user preference descriptors). Another advantage of MPEG-7 is that it offers a system's part that allows coding of descriptions (including compression) for streaming and for associating parts of MPEG-7 descriptions to media units, which they describe. In fact,

* Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1095–1102, 2008.
© Springer-Verlag Berlin Heidelberg 2008

in [1] L. Rutledge and P. Schmitz proved the need of a media in format MPEG-7 to improve media fragment integration in Web document can be done until now only with textual document as HTML or XML document. TV anytime with their vision of future digital TV that offers the opportunity to provide value-added interactive services, has also claimed that MPEG-7 collection of descriptors and description schemes is able to fulfill the metadata requirements for TV anytime.

As more and more audiovisual information becomes available from many sources around the world and people would like to use it for various purposes, there are many standards including but not limited to RDF Site Summary (RRS), SMPTE Metadata Dictionary, EBU P/Meta, TV-Anytime. The Resource Description Framework (RDF) developed by the World Wide Web Consortium (W3C) – provides the foundation for meta-data interoperability across different resource description communities. Dublin Core (DC) is an RDF-based standard that represents a meta-data element set intended to facilitate the discovery of electronic resources. Dublin Core is currently used as a meta-data standard in many TV archives. Hunter et al. showed in [2] that it is possible to describe both the structure and fine- grained details of video content by using the Dublin Core elements plus qualifiers. TV-Anytime specifications provides very little low level information but provides a large panel of semantically higher level information (e.g., title, synopsis, genre, awards, credits, release information etc.). Besides, since TV-Anytime is initially dedicated to TV services, it handles the segmentation and the repeatable broadcast of multimedia resources or programs. It is also interesting to note that TV-Anytime uses a subset of the MPEG-7 standard for low-level information.

These standards provide basic features or basic tools for developing metadata applications. For specific applications we have to refine these standards. Therefore, in this paper, we focused on designing and developing this system to integrate of multimedia metadata from heterogeneous sources, which can enables the collection, analysis and integration of multimedia metadata semantic description from some different kinds of services (UCC, IPTV, VOD and Digital TV et al.). This paper is organized as follows: section 2 discusses related work; section 3 shows our Multimedia Metadata Management System architecture and detailed system analyses; section 4 introduces the annotation technology; section 5 shows the implement and results, finally, conclusion and future work.

2 Related Work

Multimedia Database Management Systems (MMDBMSs) are the technology for content management, storage and streaming. Furthermore, a MMDBMS must be able to handle diverse kinds of data and to store sufficient information to integrate the media stored within. In addition, to be useful a database has to provide sophisticated mechanisms for querying, processing, retrieving, inserting, deleting and updating data. SQL/MM, the Multimedia Extension to SQL developed by the SC 39 of ISO/IEC supports multimedia database storage and content-based search in a standardized way [3]. The content-based multimedia retrieval and search have power to improve services and systems in digital libraries, education, authoring multimedia, museums, art galleries, radio and TV broadcasting, e-commerce, surveillance, forensics, entertainment, so on [4].

MPEG-7 and TV-Anytime are established standards for defining multimedia contents. They answer to various needs and thus, they are complementary on several levels. MPEG-7 [4] supplies a set of tools to describe a wide range of multimedia contents. Although MPEG-7 [5] covers a wide range of abstraction level, the MPEG-7 strength is inherent in the description of low-level information such as texture, color, movement and position information and also in the specifications of description schemes. In contrast, TV-Anytime with their vision of future digital TV that offers the opportunity to provide value-added interactive services, has also claimed that MPEG-7 collection of descriptors and the TV-Anytime specifications provides very little low level information but provides a large panel of semantically higher level information (e.g., title, synopsis, genre, awards, credits, release information etc.). Besides, since TV-Anytime is initially dedicated to TV services, it handles the segmentation and the repeatable broadcast of multimedia resources or programs. It is also interesting to note that TV-Anytime uses a subset of the MPEG-7 standard for low-level information.

In this paper, our system also focuses on indexing the multimedia metadata according to its content, to allow users browse the multimedia effectively. Automatic video indexing technique is essential to reduce the required time and human work in indexing large video data. Some recent studies have already used specific-domain approach in developing solutions for content video description schemes [6,7] and XML document retrieval [8]. Zhou et al. [9] use a similar approach to use a rule-based indexing system for basketball video. Rui et al. [10] used an experiment with baseball games to test the accuracy of detecting the combination of two highly correlated audio aspects, which are highly correlated with exciting events: commentator's excited speech and baseball pitch and hit. C. Tsinaraki et al. [11] proposed a framework and the software infrastructure that developed for the support of ontology-based semantic indexing and retrieval of content following the MPEG-7 standard specifications for metadata descriptions. The framework is also applicable for content metadata which have been structured according to the TV-Anytime standard and the same ontology can be used to guide the TV-Anytime content and user profile metadata derivations and retrieval processes.

3 System Design

3.1 Overview

In this section, we will describe a new Multimedia Metadata Management System, which can support the integration of multimedia metadata from heterogeneous sources. Also, enables the collection, analysis and integration of multimedia metadata semantic description from some different kinds of services.

Generally, when we search images or videos from several different kinds of services, these specific objects will be not easily to be found over all kinds of services by internet. That's for these services are always based on some different standards and these standards are not easily to be shared from each other. MPEG-7 improve media fragment integration in Web document, and the media fragment integration in

Web document can be done until now only with textual document as HTML or XML document. TV Anytime with their vision of future digital TV that offers the opportunity to provide value-added interactive services, has also claimed that MPEG-7 collection of descriptors and description schemes is able to fulfill the metadata requirements for TV Anytime. How to integrate multimedia metadata from heterogeneous sources is our basic review of our system. So, in the next section, we will explain our system's architecture in detail.

3.2 System Design

In our system, we propose this metadata integration management system composed of four main units: Metadata Acquisition unit, Metadata Analyzer unit, Metadata Mapping unit, and Metadata storage unit. (As shows in Figure 1)

Users research for Multimedia data throughout internet sharing by transferring XML documents, content-based metadata information can be managed by Metadata Acquisition Manager. Then, these acquired XML metadata documents will be filtered and reparsed by the Metadata Analyzer; and next transformed into a uniform metadata semantics description by Metadata Mapping Manager. Finally, these segments metadata information are stored into the Metadata repository for further processing and search/retrieval.

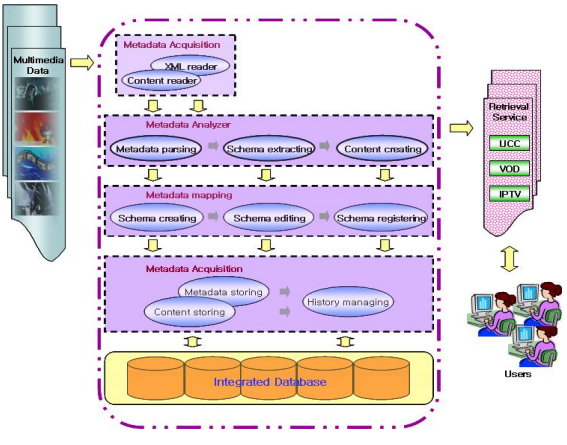


Fig. 1. System Design

3.3 Module Function

Metadata Acquisition

Metadata Acquisition customized generator for input multimedia document. It used the TCP/IP protocol to acquire multimedia document. The Metadata Acquisition algorithm is shown in Figure 2.

```
Algorithm FileAcquisition (meta, content)
Input : meta (metadata file), content (multimedia file)
Output : iMeta (metadata instance), iContent (multimedia instance)
Begin
    if(File Type is XML) then
      imeta = create text file
    else
      icontent = create binary file
    end if
    while(EOF of meta or content)
        imeta = readFile(meta)
        icontent = readFile(content)
    end while
    closeFile(imeta, icontent)
    return imeta, icontent
end
```

Fig. 2. Metadata Acquisition Algorithm

Metadata Analyzer

Metadata Analyzer customized generator for generated instance by Metadata Acquisition. Metadata Parsing used to filter and reparse content-based metadata which are identified in Metadata Acquisition manager. Schema Extracting used to classify the complex schemas for every different standard. Content Creating summarize these different content-based metadata semantics descriptions and create a new content table which contains the common descriptions.

Metadata Mapping

Metadata Mapping manager model as a unique model that compares and redefines different types of standard schemas into a common identified one to satisfy requirement applications from also different services immediately. This model not only can reparse these metadata semantics descriptions into some common integrate schemas, but also design a annotation table for supporting some specific semantics descriptions which are not contained into common integrate schemas. We can get more accurate request from all shared services in internet; also, we can capture some specific metadata information by indexing annotation table with each single service. That's why our system can support the integration of multimedia metadata from heterogeneous resource internet sharing and users can get a high-quality request from any kind of service immediately. The table of Mapping Schema is presented in following part.

Table 1. Metadata Mapping Schema

Korean Name	English Name	Description
	attributeName	Define the name of schema attributes in this system.
	attributeDescription	Describe the schema attribute.
	matchedAttribute	Define the same attribute name for the same content in this system, although the original attribute names are different.
	matchedAttributeValue	Matching the same attribute values.
	attributestandard	Matching the same standard names.
	attributeIndex	Index some different attributes for the same contents.

Metadata Storage
Metadata Storage customized generator for integrated data by previous managers. The Metadata Storage algorithm is shown in Figure 3.

```
Algorithm ExtractSchema (mappedListSet, icontent)
Input : mappedListSet (mapped schema list instance),
        icontent (content instance)
Output : boolean (outcome)
Begin
    dbCon = connect database
    temp = find CRID of mappedListSet in metadata table
    if(exsists CRID in metadata table)
        historicalrecord = find record having CRID of mappedListSet
                           in metadata table
    end if
    insertMetadata (mappedListSet)
    insertContent (icontent)
    return true
end
```

Fig. 3. Metadata Storage Algorithm

4 Annotation

While designing this architecture we have considered its use in two main contexts. These applications cover a range of requirements in both personal and professional management of multimedia information: One is the management of personal multimedia collections of data that includes the archiving and the retrieval of specific items under particular semantic conditions; and the other is the management of professional multimedia data within a network to share multimedia resources and related semantic information, where ownership and authorization rights should be taken into account. Therefore, we should consider that how to store, organize and retrieve distributed multimedia resources; how to manage algorithms for information processing; how to add semantic annotations; and how to access, protect and/or share information.

5 Implementation

The implement environment scenario as follows: First, capture the transferring XML documents from internet or external transforms. Then we can extract the common attributes and create a new XML document after Tree-view analyses. Next, we set 7 kinds of attributes (Title, Actor, Description, Producer, Support, Genre, and Date) to test both compatibility and accuracy for combining three various standards (ADI, TV-Anytime and MEPG-7) after Metadata Mapping Manager processing.

The Figure 4 shows the original datasets which integrated XML documents have been parsed by Metadata Mapping Manager processing. The implementation results appear that our system supports distributed services resource for a high sharing and searching throughout internet, although these services are based on unshared standards. Our approaches also help the system to reduce the complexity of various standards description data.

Table 2. Implementation Environment

CPU	Intel(R) Core(TM)2 6600 @ 2.40GHz
RAM	1GB RAM
System	Booyo Linux
Language	JDK_1.6
Program Tool	eclipse 3.2
XML Parser	SAX, DOM
Database	My-SQL_5.0
Database Interface	JDBC_5.0

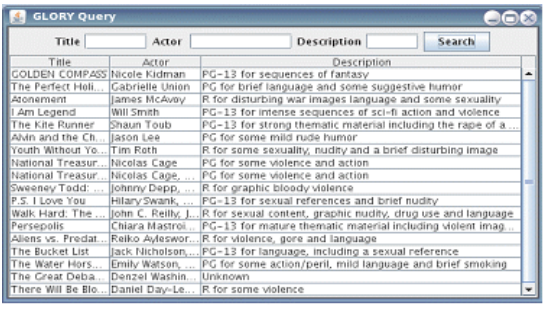


Fig. 4. Original Datasets

6 Conclusions

In this paper, we presented a new integration Multimedia Metadata Management System, which can support the integration of multimedia metadata from heterogeneous sources. Actually, the restriction models (Metadata Mapping Manager) following the common semantics description schemas can help the system to reduce the complexity of unshared standards description data.

However, it is only tested three services (ADI, TV-Anytime and MEPG-7) in which we have only experimented with the structure of media content. Further experiment works focus on more complex content-based search over more various services especially for some emerging standards. Moreover, we are also working on enhancing the Web-based query interface of our system for multimedia query specification support.

Acknowledgments. This research was supported by a grant (#07KLSGC02) from Cutting-edge Urban Development - Korean Land Spatialization Research Project funded by Ministry of Construction & Transportation of Korean government and Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (The Regional Research Universities Program/Chungbuk BIT Research-Oriented University Consortium).

References

1. Lloyd, R., Patrick, S.: Improving Media Fragment Integration in Emerging Web Formats. In: 8th International Conference on Multimedia Modeling (MMM 2001), pp. 147–166. CWI Press, Amsterdam (2001)
2. Jane, H.: A Proposal for the Integration of Dublin Core and MPEG-7. In: ISO/IEC/JTC1/SC29/WG11 M650054th MPEG Meeting. La Baule (2000)
3. Harald, K.: Distributed multimedia Database Systems supported by MPEG-7 and MPEG-21. CRC Press, Boca Raton (2003)
4. Manjunath, B.S., Salembier, P., Sikora, T.: Introduction to MPEG-7: Multimedia Content Description Interface, pp. 335–361. John Wiley & Sons Inc. Press, Chichester (2002)
5. ISO MPEG-7, part 5-Multimedia Description Schemes. ISO/IEC/JTC1/SC29/WG11/N4242 (2001)
6. Olivier, A., Philippe, S.: MPEG-7 Systems: overview. In: IEEE Transactions on Circuits and Systems for Video Technology, pp. 760–764. IEEE Press, Madison (2001)
7. Silvia, P., Uma, S.: TV anytime as an application scenario for MPEG-7. In: ACM Multimedia Workshops 2000, pp. 89–92. ACM Press, Los Angeles (2000)
8. Scott, B., Don, C., Mary, F.F., Daniela, F., Jonathan, R., Jerome, S., Mugur, S.: XQuery 1.0: An XML Query Language. In: W3C Working Draft (2001)
9. Wensheng, Z., Asha, V.: Rule-based video classification system for basketball video indexing. In: ACM Multimedia Workshops 2000, pp. 213–216. ACM Press, Los Angeles (2000)
10. Yong, R., Anoop, G., Alex, A.: Automatically extracting highlights for TV baseball programs. In: ACM Multimedia 2000, pp. 105–115. ACM Press, California (2000)
11. Chrisa, T., Panagiotis, P., Fotis, K., Stavros, C.: Ontology-based Semantic Indexing for MPEG-7 and TV-Anytime Audiovisual Content. Multimedia Tools Appl. 26(3), 299–325 (2005)

Study of Image Splicing Detection

Zhen Zhang[1,2], Ying Zhou[3], Jiquan Kang[2], and Yuan Ren[2]

[1] Institute of Information Engineering, Information Engineering University, Zhengzhou 450002, China
[2] School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China
[3] College of Materials Science and Engineering, Zhengzhou University, Zhengzhou 450001, China
zhangzhen66@126.com

Abstract. With the advent of digital technology, digital image has gradually taken the place of the original analog photograph, and the forgery of digital image has become more and more easy and indiscoverable. Image splicing is a commonly used technique in image tampering. In this paper, we simply introduce the definition of image splicing and some methods of image splicing detection, mainly including the detection based on steganalysis model, the detection based on Hilbert-Huang transform (HHT) and moments of characteristic functions (CF) with wavelet decomposition. We focus on discussing our proposed approach based on image quality metrics (IQMs) and moment features. Especially we analyze the model creation and the extraction of features in digital image. In addition, we compare these approaches and analyze the future works of digital image forensics.

Keywords: Image splicing detection, Digital image forensics, Passive-blind forensics, Image feature, Image quality metrics (IQMs).

1 Introduction

Unlike text which is called conventional media, images provide a vivid and natural communication media for human, as human often need no special training to understand the image's content. Therefore, photograph soon became the most popular and convictive media just after the photograph technology's invention. Though a lot of new media's flourish in recent decades, image still plays an important role in our life. With the advent of digital times, the digital data has gradually taken the place of the original analog data, and in recent years, due to the advent of high-performance commodity hardware and improved human-computer interfaces, it has become relatively easy to temper images. At the beginning, the alternation is just enhance the image's performance, but then many people started to alter the image's content, even to gain their ends by these illegal and immorality methods. Based on the above reasons, it's valuable to develop some credible methods to detect whether a digital image is tempered, and as the most fundamental and common mean of image forgery, image splicing is a valuable topic to study.

In section 2, we introduce some methods of digital image tampering and give a definition of the spliced image. In section 3, we sum up some methods of image splicing

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1103–1110, 2008.
© Springer-Verlag Berlin Heidelberg 2008

detection and propose a novel model based on image quality metrics (IQMs) and moment features. In section 4, we conclude the paper and analyze the future works of digital image splicing detection.

2 Image Splicing

There are many methods of digital image tampering, such as image compositing, image morphing, image re-touching, image enhancement, computer generating, and image rebroadcast. Image compositing is the commonest method, and image splicing is a basic operation in it.

Image splicing is a technology of image compositing by combining image fragments from the same or different images without further postprocessing such as smoothing of boundaries among different fragments. The steps of image splicing are shown in Fig.1.

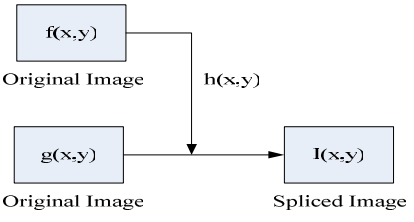


Fig. 1. The steps of image splicing. $f(x, y)$ and $g(x, y)$ are original images. $h(x, y)$ is a part of $f(x, y)$ which is insert into $g(x, y)$ and generate spliced image $I(x, y)$. Perhaps $f(x, y)$ and $g(x, y)$ is the same image.

Many people have done a lot of work on image splicing detection. There are many methods which have been proposed including the detection based on the consistency of image content and the detection based on statistical mode. We will make detailed descriptions as follows.

3 The Detection of Image Splicing

3.1 The Detection of Image Splicing Based on Hilbert-Huang Transform and Moments of Characteristic Functions with Wavelet Decomposition

This approach was presented by Dongdong Fu, Yun Q. Shi, and Wei Su in [1]. In this paper, the image splicing detection problem is tackled as a two-class classification problem under the pattern recognition framework shown Fig.2. Considering the high non-linearity and non-stationarity nature of image splicing operation, Hilbert-Huang Transform (HHT) is utilized to generate features for classification. Furthermore, a statistical natural image model based on moments of characteristic functions (CF) with wavelet decomposition is used to distinguish the spliced images from the authentic images. Support vector machine (SVM) as the classifier is used.

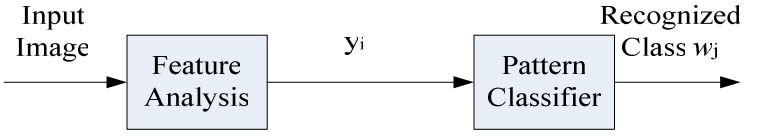


Fig. 2. Pattern recognition framework. $y_i = R^n$ is a feature vector and its associated class is $w_j \in \{-1,1\}$. w_j = 1 if y_i is spliced image, otherwise w_j = -1. The classifier learns the mapping $y_i \rightarrow w_j$ and produces a decision function. Classifiers can use Bayes, support vector machine and neural network.

The detail procedure for extracting HHT features is shown in Fig.3.

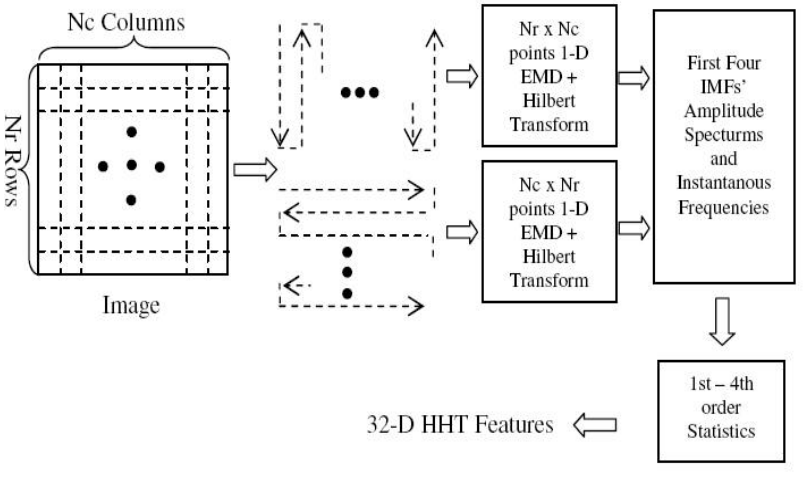


Fig. 3. The HHT feature extraction procedure. EMD is the abbreviation of empirical mode decomposition. IMFs are the abbreviation of intrinsic mode functions. After HHT, there are totally 32-D features generated.

In addition to the features extracted by using HHT, the approach also proposes a natural image model to capture the differences between authentic image and spliced image which is introduced by splicing operation. The definition of model features extracted by the statistical natural image model based on moments of characteristic functions with wavelet decomposition is listed as follows.

$$M_n = \sum_{j=1}^{(N/2)} f_j^n \left|H\left(f_j\right)\right| \Big/ \sum_{j=1}^{(N/2)} \left|H\left(f_j\right)\right| \tag{1}$$

Where $H\left(f_j\right)$ is the CF component at frequency f_j, N is the total number of points in the horizontal axis of the histogram. There are 78-D features generated.

This scheme can achieve a promising detection accuracy of 80.15%.

3.2 The Detection Based on Steganalysis Model

Steganalysis is the science of detection the very existence of the hidden message in a given media. The tasks include detection of existence of hidden message, estimation of message length, message extraction/destroying, etc. Both steganalysis and detection of image splicing belong to forensics and use digital image as important carrier. So, it is feasible using steganalysis model to detect spliced image.

Chen proposed a steganalysis model which can be used on image splicing in [2]. This scheme utilizing first three statistical moments of characteristic functions of the test image, its prediction-error image, and all of their wavelet subbands is proposed in [3], where a 78-D feature vector is used as the image model. In addition, the model present a steganalyzer which combines statistical moments of 1-D and 2-D characteristic functions extracted from the image pixel 2-D array and the multi-size block discrete cosine transform (MBDCT) 2-D arrays. Chen's model is as follows:

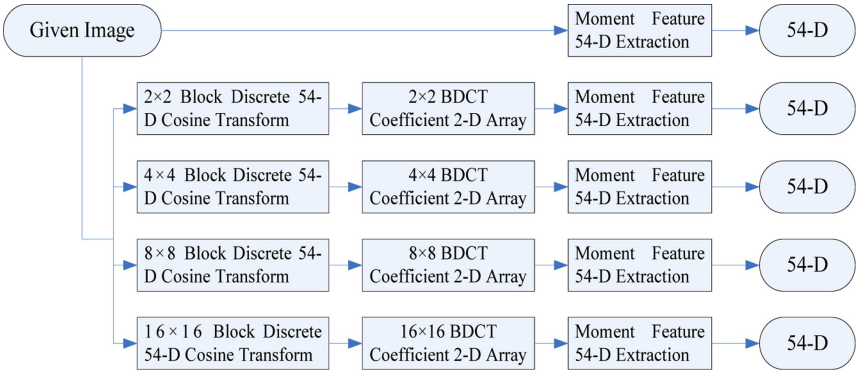


Fig. 4. Chen's steganalysis model

It performs best using this steganalysis model to detect spliced image and achieves a promising detection accuracy of 87.07%.

3.3 The Detection Based on Image Quality Metrics and Moment Based Features

In this section, we present a novel approach to passive detection of image splicing based on image quality metrics and moment features. Our experimental works have demonstrated that this splicing detection scheme performs effectively. Shown in Fig. 5, image splicing detection problem can be regarded as a two-class classification problem under the pattern recognition framework. The features are extracted by image quality metrics (IQMs) and moments of characteristic functions of wavelet subbands.

Image Quality Metrics. There is a wealth of research on subjective and/or objective image quality measures to reliably predict either perceived quality across different scenes and distortion types or to predict algorithmic performance. Avcibas has presented collectively a set of image quality measures which are sensitive and discriminative to compression, watermarking, blurring and noise distortions in [4] and [5]. We chose seven

most sensitive metrics to detect image splicing operation as features. There are that mean absolute error D1 and mean square error D2 based on statistical difference between images and filter version of its, correlation measures based on statistical difference between images and filter version of its: C_4, D4, D5, spectral measures based on statistical difference between RGB images and filter version of its: S_1, human visual system measures based on statistical difference between RGB images and filter version of its: H_1.

We divide a test image into N region, then, calculate image quality metrics to the N regions which will obtain 7×N features.

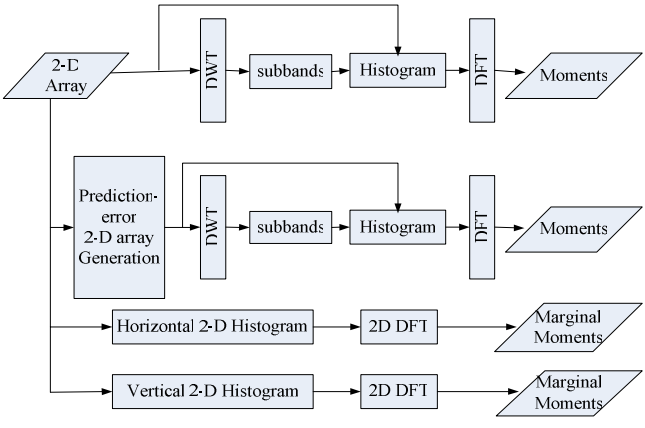


Fig. 5. Moment extraction procedure

Moment Based Features. Moment extraction procedure is shown in Fig.5.

2-D Array may be image 2-D array, 2×2, 4×4, 8×8 block discrete cosine transform (BDCT) to image array.

A prediction-error 2-D array is used to reduce the influence caused by diversity of the image content and to simultaneously enhance the statistical artifacts introduced by splicing. The prediction context is shown in Fig.6.

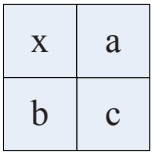


Fig. 6. Prediction context

We need to predict the value of x using the values of its neighbors a, b, and c. The prediction 2-D can be given by:

$$\bar{x} = sign(x) \cdot \{|a| + |b| - |c|\} \tag{2}$$

And prediction-error 2-D array can be expressed by:

$$\Delta x = x - \bar{x} = x - sign(x) \cdot \{|a| + |b| - |c|\} \tag{3}$$

The Model of All Features Extraction. End of the above two procedures, we can obtain the model of all features extraction shown in Fig.7.

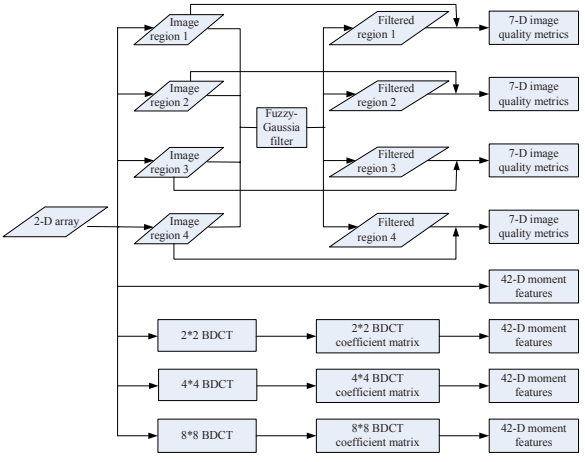


Fig. 7. The features extraction procedure

Thus, we can extract 196-D features including 4×7=28-D IQMs and 4×42=168-D moment features. The average of accuracy rate can achieve 86.70% using this model.

Table 1. Comparing between Chen' model, Dongdong Fu' model and our model

	TN	TP	Accuracy
Chen's Steganalysis Model (270–D)	86.32%	87.83%	87.07%
HHT + Model (110–D)	80.25%	80.03%	80.15%
IQMs + Model (196–D)	90.51%	87.14%	88.80%

4 Conclusions

Three models have got very plummy effect. Details are shown in Table 1.

Above models use the dataset in [6]. There are totally 933 authentic images and 912 spliced images in this dataset. In the classification process, Dongdong Fu' model randomly selected 5/6 of total 933 authentic images and 5/6 of total 912 spliced images for training and the remaining 1/6 of the authentic images and spliced images for testing the trained classifier. In our experiment we randomly selected 85% of

spliced images and authentic images to train a SVM classifier. Then the remaining 15% of spliced images and authentic images to test the trained classifier. It is easy to know that our proposed model performs best.

5 Future Works

The detection of image splicing has got considerable achievement; still, it has left much to be desired. Some of future works are as follows:

1. Since detection rates are still lower, the research for high detection rates is necessary.
2. Researching faster algorithms.
3. Building large and realistic image database.
4. To make the forgery even harder to detect, one can use the feathered crop or the retouch tool to further mask any traces of the splicing segments. So tampering other than splicing need to be investigated.

Overcoming these challenges requires the development of several novel methodologies and thorough evaluation of their limitations under more general and practical settings. This can be achieved in collaboration with forensics experts and through their continuous feedback on the developed methods. The research effort in the field is progressing well in these directions.

References

1. Fu, D.D., Shi, Y.Q., Su, W.: Detection of Image Splicing Based on Hilbert-Huang Transform and Moments of Characteristic Functions with Wavelet Decomposition approach, http://www.springerlink.com/index/fp36p770x04r9301.pdf
2. Chen, C., Shi, Y.Q.: Steganalyzing Texture Images. IEEE ICIP (2007)
3. Shi, Y.Q., Xuan, G., Zou, D., Gao, J., Yang, C., Zhang, Z., Chai, P., Chen, W., Chen, C.: Steganalysis Based on Moments of Characteristic Functions using Wavelet Decomposition, Prediction-error image, and Neural Network. In: IEEE ICME (2005)
4. Avcibas, I., Sankur, B., Sayood, K.: Statistical Analysis of Image Quality Measures. Journal of Electronic Imaging 11(4), 206–223 (2002)
5. Avcibas, I., Memon, N., Sankure, B.: Steganalysis using Image Quality Metrics. IEEE Trans. Image Processing 12(2), 221–229 (2003)
6. Columbia Image Splicing Detection Evaluation Dataset, DVMM, Columbia University, http://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet/AuthSplicedDataSet.htm
7. Alin, C.P., Hany, F.: Exposing Digital Forgeries by Detecting Duplicated Image Regions, http://www.cs.dartmouth.edfarid/pub1ications / tr04.pdf
8. Chang, C.C., Lin, C.J.: LIBSVM – a library for support vector machines (2008), http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html
9. Jessica, F., David, S., Jan, L.: Detection of Copy-Move Forgery in Digital Images. In: Proceedings of Digital Forensic Research Workshop, Cleveland, OH (2003)
10. Johnson, M.K., Farid, H.: Exposing Digital Forgeries by Detecting Inconsistencies in Lighting. In: ACM Multimedia and Security Workshop, New York, NY (2005)

11. Wang, Y.X.: Talking about the Original Examination of the Static Picture. Journal of Sichuan Police College 18(1) (February 2006)
12. Ng, T.T., Chang, S.F.: Blind Detection of Photomontage Using Higher Order Statistics. In: Proceedings of the 2004 International Symposium, V-688-V-69. IEEE, Vancouver (2004)
13. Ng, T.T., Chang, S.: A Model for Image Splicing. Image Processing, 2004. In: ICIP 2004. International Conference, pp. 1169–1172. IEEE, Singapore (2004)
14. Kovesi, P.: Image Features from Phase Congruency. Journal of Computer Vision Research 1(3), 1–26 (1999)
15. Chen, W., Shi, Y.Q., Su, W.: Image Splicing Detection using 2-D Phase Congruency and Statistical Moments of Characteristic Function. In: SPIE Electronic Imaging: Security, Steganography, and Watermarking of Multimedia Contents, San Jose, CA, USA (January 2007)
16. Shi, Y.Q., Chen, C., Chen, W.: A Natural Image Model Approach to Splicing Detection. ACM MM & Security (2007)

Research on Real-Time Sequential Decision of Texas Star Assembly

Ying Zhou[1], Zhen Zhang[2,3], and Chenyong Liu[4]

[1] College of Materials Science and Engineering, Zhengzhou University, Zhengzhou, 450001, Henan, China
[2] College of Information Engineering, Information Engineering University, Zhengzhou, 450002, Henan, China
[3] School of Electric Engineering, Zhengzhou University, Zhengzhou, 450001, Henan, China
[4] Department of Battle Command, Air Defence Command College, Zhengzhou, 450052, Henan, China
zhouying@zzu.edu.cn

Abstract. Sequential decision processes can be described as Markov processes. Because it is difficult to solve Markov processes, the real-time sequential decision processes based on VR simulation, which avoids to calculate the state transition probability matrix, is discussed. A new mechanism of driving directly VRML scene by simulator which is developed by VRMLScript is put forward. The mechanism solves the thorny problem of simulation control, such as scene glint and logic error of simulation etc. The real-time sequential decision system of texas star (a component of a passenger plane) assembly is developed .It is proved that the expected goal of the real-time sequential decision of texas star assembly can be approaching gradually.

Keywords: Real-time sequential decision; Simulation; VRMLScript; Virtual reality.

1 Introduction

Virtual Reality (VR) provides a powerful tool to interact with computers and complicated graphical data. The Virtual Reality Simulation system combines VR technology and discrete event system simulation, which may make simulation more intuitive and creditable because users can interact with and immerse into the virtual world. Under the environment of virtual reality simulation, decision makers can see clearly the essence of a very complicated decision problem from different view points and angles during its implementation, decisions can be made efficiently in a real-time fashion. As he or she finds the difference between current performance measure caused by initial decisions and expected results of the model or an important rare event occurs, then the simulation process can be suspended and decision parameters of the model can be modified, after doing so, the simulation process is resumed .When the real-time sequential decision process is performed, the expected goal of the problem will be approaching gradually.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1111–1117, 2008.
© Springer-Verlag Berlin Heidelberg 2008

VR simulation technology based on Web is being regarded widely now. After VRML97 specification was published by SGI Corporation in Aug, 1996, VRML has become a perfect language for implementing VR simulation over Internet. At present, the VR simulation mechanism based on VRML is that VRML scene is driven directly by general simulation software or simulation software package. However, when computer system has a lower configuration, the thorny problem of simulation control, such as scene glint and logic error of simulation etc, occurs because of a large numbers of data exchange between general simulation software or simulation software package and VRML scene. In order to solve the problem, a new mechanism under which VRML scene is driven directly by the simulator which is developed by VRMLScript and is of the function of real-time sequential decision is put forward. The real-time sequential decision system of texas star (a component of a passenger plane) assembly is developed using the new mechanism, and it is proved that the new mechanism is effective.

2 VR Simulation System

2.1 VRML on Web for Simulation

VRML stands for "Virtual Reality Modeling Language". It is a file format for describing interactive 3D objects and virtual worlds over internet. The file format is a subset of Open Inventor which was developed by Silicon Graphics Company [1,2]. To work with VRML, a virtual world player is required, which can explain the VRML file and use it to build a 3D VR world. The player is provided as plugins of the Internet browser such as Netscape's Navigator and the Microsoft's IE. We choose Cosmo Player because of its good performance [3].

In VRML file, "node" is the most important and fundamental syntax unit. It is an abstract of various real-world object and concept. Nodes contain fields and events, messages may be sent between nodes along route. There are 54 standard nodes in VRML97.VR scene is built by using these nodes.

VRML allows us to extend its ability. There are some programming languages that we should use with VRML, and these programming languages can code a specific simulator which can complete the common simulation work and also drive VR objects at the same time.

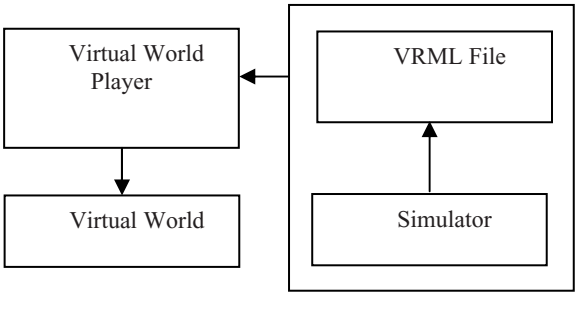


Fig. 1. Relationship of VWP, VRML & Simulator

The relationship of VRML, Simulator and virtual world player can be explained by Fig 1. The virtual world player reads and explains the VRML file and build a 3D virtual world. The VRML file just tells the virtual world player how to build the predefined virtual world, the Simulator completes the simulation work and drives the virtual world run dynamically.

2.2 Simulator Based on VRMLScript

The simulator is not only the simulation engine which controls the running of simulation, gathers simulation statistical data, but also a VR driver which can dynamically add/remove the simulation entities and control their movement, etc.

There are two ways to create a VR simulation system. The usual way is using Java API to create a simulator and VRML objects are driven by the Java based simulator. However the problem of simulation control, such as scene glint and logic error of simulation etc, is difficult to be solved [5]. We use VRMLScript to code the simulator and Java is mainly used to create the control interface. This gives possibility to user who could control the VRML- based VR simulation system conveniently, so the control problem can be solved. This mechanism is suitable to the real-time sequential decision process.

There is a node called Script in VRML. It is an interface between VRML and other programming language such as VRMLScript and Java. It's field "url" gives the reference to the specific simulator program. For example

```
Script{
      url "vrmlscript: function init(){…} …"
}.
```

To access the node in VRML from code the field direct Output must be set to TRUE. In the Script node we can also add some user-defined fields to store extra data. The use of VRMLScript see the reference 1. Next, we discuss mainly algorithm according to character of VRMLScript.

- Simulation clock

Simulation clock drives the simulation run. It is very important for simulation. The Time Sensor node in VRML can trigger events continuously after a predefined period of time, so it can be used as the simulation clock. Start and stop of the simulation clock may be implemented by the field enabled, startTime and stopTime. When startTime>stopTime and enabled=TRUE, the simulation clock is started. When enabled=FALSE, the simulation clock is stopped.

- Dynamic behavior of entities

How to add/remove entities and control their movement dynamically in the VR system is the key problem for the animation of VR Simulation. A system object Browser can get the goal. Browser object can be used directly in VrmlScript. We can use the method create VrmlFromString (String vrmlSyntax) to add an entity, use the methods addRoute(, , ,) and deleteRoute(, , ,) to control the movement of the entity, use field removeChildren in VRML nodes to remove an entity. Detailed information about Browser object please refers to reference 2 and 3.

- Pausing and restarting simulation

In order to support real-time sequential decision, the simulator should have the function of pausing simulation, modifying parameters of VR simulation system and restarting simulation. Because VRMLScrit dose not have the function of accepting user's input, Java Applet is used to carry out communication between user and VRMLScript.

Pausing simulation and restarting simulation can be carried out by controlling field enabled of TimeSensor node .First, get the instance of TimeSensor node by method getNode() of browser object, then get the instance of field enabled by method getEventIn(), finally, set value of field enabled by method setValue(). The codes are as follows.

```
// get the instance of TimeSensor node which stands for
the simulation clock
sim_control_time=browser.getNode("sim_control_time");
// restart simulation
EventInSFBool ev=(EventInSFBool);
sim_control_time.getEventIn("enabled");
ev.setValue(true);// Pause simulation
EventInSFBool ev=(EventInSFBool);
sim_control_time.getEventIn("enabled");
ev.setValue(false);
```

The method of modifying parameters of VR simulation system is similar to the method above. First, get simulation parameters of user input in Javaapplet, and then send the simulation parameters to simulator by method setValue().

3 Real-Time Sequential Decision Based on VR Simulation

Sequential decision processes can be described as Markov processes. It is difficult to solve Markov processes because we have not found a good method to calculate the state transition probability matrix [4,5]. Sequential decision based on VR simulation, which avoids to calculate the state transition probability matrix, is sequential decision under supported by VR simulation system.

Decision points (Where), levels of decisions at each decision point (What), decision time (When) and Number of decisions (Number) are fundamental parameters of the model of sequential decision processes based on VR simulation. We conclude this as "3W+N" mode. The decision points stand for the supervisor's preference. The levels of decisions at each decision point stand for the intensity of decisions. The decision time indicates when to make decision. The number of decisions indicates how much we will make decision.

During sequential decision processes based on VR simulation, the supervisor can pause or restart VR simulation system whenever he or she found any abnormal status was happening, then modify decision parameters, and then restart the system. We call this fashion as real-time sequential decision processes based on VR simulation. During real-time sequential decision processes based on VR simulation, interaction between the supervisor and the system is emphasized, and the decision time is determined according to status of the system by the supervisor, so the real-time sequential decision is more validated.

4 Real-Time Sequential Decision of Texas Star Assembly

As a case study, we formulate a model for assembling fuselage of Boeing 737(named Texas Star). Fig 2 shows the real-time sequential decision system of texas star assembly. The Texas star is composed by 6 kinds of beams, which are created to enter the system according to negative exponential distributions with different parameters form 0 to 1.5, other parts needed by the assembly line are assumed having enough inventory in the workshop. The assembly line is consist of 7 machines which are coded as 9801L, 29801L, 9801R, 29801R, 9800, 9800b and 9805.The machine 9800b is backup machine. The assembly process includes operations of initial assembly, Texas star riveting, supplementary riveting, precise processing of the star etc, all the processing times are assumed to be triangular distributions with different parameters.

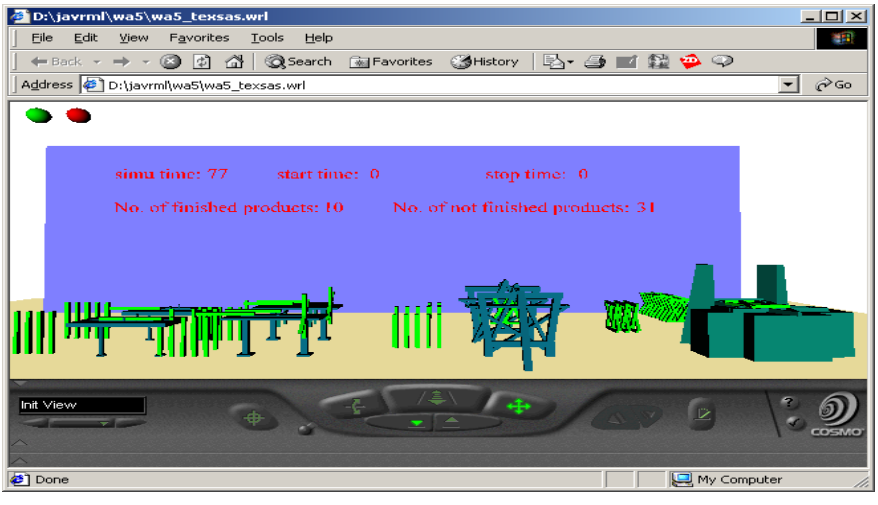


Fig. 2. Real-time sequential decision system of texas star assembly

During the assembly process, the supervisor can browse on the screen and enter into the VR scene at any point, once he or she found any abnormal status was happening, a real time decision may be taken to modify arrival rate when level of WIP in the process is higher, or to start the backup machine, when a main machine is failed or the production capacity is not enough. In this assembly line, every machine is a decision points .The levels of decisions at each decision point are infinite because of continuity of the arrival rate.

The purpose of this decision procedure is to make the assembly process with lower level of WIP by adjusting the arrival rate of beams. It is acceptable and economic that the level of WIP is between 1 and 4 .Now we do simulation during a long time .The real-time sequential decision processes is made in term of rules as follows:

- When the average level of WIP is greater than 4, then decrease the arrival rate of beams.
- When the average level of WIP is less than 1, then increase the arrival rate of beams.

Table 1 shows the decision processes. When the supervisor found the average levels of WIP of left main beam and fore beam were greater than 4, the simulation was paused, and the arrival rate of left main beam was decreased to 0.4 and that of fore beam was decreased to 0.3. When back beam's average level of WIP was greater than 4, the simulation was paused, and the arrival rate of back beam was decreased to 0.4 .The average levels of all beams are greater than 1 and less than 4 during a long time since then, so the real-time sequential decision processes based on VR simulation is over. The arrival rate of beams at present may be used during texas star assembly.

Table 1. The real-time sequential decision processes of texas star assembly

Beams	Left main beam	Right main beam	Left strengthen part	Right strengthen part	Fore beam	Back beam
Initial value of the arrival rate of beams	1.1	0.8	0.5	0.5	1.3	0.6
Average levels of WIP at first pausing time	4	2.5	2	1.8	6.8	1.7
Adjustment	0.4	0.8	0.5	0.5	0.3	0.6
Average levels of WIP at second pausing time	1.3	2.7	2	1.9	3.7	4.6
Adjustment	0.4	0.8	0.5	0.5	0.3	0.4
Steady average levels of WIP	1.8	2.7	2.1	1.9	3.7	3.8

Besides, the VR simulation system also can be used to train the decision makers who may have different preference, once they have made a sequence of decisions, the simulation results can show the performance measures due to their decisions, so the quality of the decision maker can be evaluated.

5 Conclusion

The real-time sequential decision processes based on VR simulation can solve the problem of complicated system, deal with uncertainty information by the interaction between the supervisor and the VR simulation system. It is used not only in production management, but also in military affairs, communications and transportation etc.

References

1. Virtual Reality Modeling Language Specification, http://www.vrml.org/Specifications/VRML97
2. Cosmo Player 2.1 FAQ for Windows 95 and NT, http://cosmosoftware.com/products/player/developer
3. Wang, F., Feng, Y.C., Wei, Y.S.: Virtual Reality Simulation Mechanism on WWW. In: Proceeding of AreoScience 2000 of SPIE, pp. 123–127 (2000)
4. Peter, R., Jonathan, L.: Bio-energy with Carbon Storage (BECS): A Sequential Decision Approach to the Threat of Abrupt Climate Change, Energy, vol. 30(14), pp. 2654–2671 (2005)
5. Zhong, L., Tong, M.A., Zhong, W., Zhang, S.Y.: Sequential Maneuvering Decisions Based on Multi-stage Influence Diagram in Air Combat. Journal of Systems Engineering and Electronics 18(3), 551–555 (2007)

A New Variable-Step LMS Algorithm Based on the Convergence Ratio of Mean-Square Error(MSE)

Hong Wan, Guangting Li, Xianming Wang, and Chai Jing

School of Electrical Engineering, Zhengzhou University, Zhengzhou 450001, China
wanhong@zzu.edu.cn

Abstract. A new variable step-size(VSS) LMS adaptive algorithm based on the convergence ratio of MSE and the correlation between reference signal and output error is proposed in the paper. Theory analyzing and simulation results prove that the new algorithm improves the convergent speed of general LMS algorithm and optimizes the trace ability of time-varying system and stable state maladjustment; and comparing to the standard LMS algorithm, the increase of computation is finite.

Keywords: adaptive filter, LMS, variable step, MSE.

1 Introduction

The LMS algorithm[1,2,3] proposed by Widrow and Hoff in 1960, is widely applied to many fields leading to its simplicity, less computation and the ease of implementation in term of hardware, such as adaptive control, radar, echo cancellation and system identification. Fig.1 shows the basic structure of the adaptive filter.

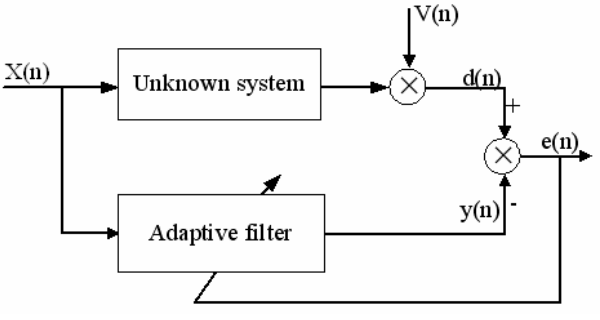


Fig. 1. the basic structure of the Adaptive filter

The standard LMS algorithm recursion is given by

$$y(n) = X^T(n)\, W(n) \tag{1}$$

$$e(n) = d(n) - y(n) \tag{2}$$

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1118–1123, 2008.
© Springer-Verlag Berlin Heidelberg 2008

$$W(n+1) = W(n)+2\mu e(n)X(n) \tag{3}$$

Where n is the iteration number, $X(n)$ is the adaptive filter input signal vector, $W(n)$ is the vector of adaptive filter weights, $d(n)$ is the desired signal, $e(n)$ is the output error of the unknown system, $V(n)$ is interference which is a zero-mean independent and μ is the step-size. The condition of algorithm achieving convergence is $0 < \mu < 1/\lambda_{\max}$, where $\lambda_{\max}$ is the maximal eigenvalue of the input signal autocorrelation matrix.

Initial speed of convergence, tracking ability in the steady state and steady-state maladjustment error are three most important aspects to evaluate the performance of adaptive filter algorithm[4]. However, the conventional adaptive filtering algorithms using fixed step size can't satisfy the acquirements at the same time[9]. In order to achieve a fast initial convergence speed and to retain a fast tracking ability in the steady state, a relatively large step size is usually chosen in practice. On the other hand, a large step size will result in large steady-state maladjustment error and make the adaptive algorithm sensitive to interference $V(n)$,appearing on the primary input. Observing the three curves of (a),(b),(c) in Fig.2,the contradiction between convergence speed and steady-state maladjustment can be found easily.

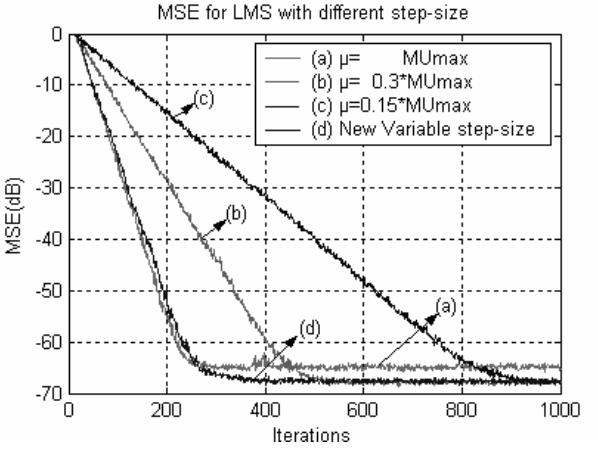


Fig. 2. The convergence speed for LMS with different step-size (MUmax=λ_{max})

In order to solve this problem, many researchers have proposed various VSS adaptive filtering algorithms. RD.Gitlin[4] presented an adaptive filtering algorithm whose step-size decreases with the increase of the iteration number n, and it can satisfy the acquirements of initial speed of convergence and steady-state maladjustment error for time invariant system at the same time, but abates for time-varying system; The algorithm proposed in reference[5],which make the step coefficient μ proportion to the output error $e(n)$, is too sensitive to interference $V(n)$; Reference [7] presented a VSS adaptive filtering algorithm (VS-NLMS) which used an estimated cross-correlation function between the input signal $X(n)$ and the error signal $e(n)$ as the VSS basis, but the calculation for cross-correlation function will takes too much time which limits its application for real-time tracking.

Considering on the advantage and disadvantage of various VSS algorithms[4~10], the principle of how to adjust step size is generalized as following:

(1) During initial convergence, a large step size should be chosen for a fast initial convergence speed.

(2) While in steady state, if the error $e(n)$ mutates, the reason for the change should be found at first. When the parameters of the unknown system vary, a large step size should be chosen , which will assure the fast tracking ability; if the change is caused by the interference $V(n)$,a small step size should be chosen for guarantying the operation of adaptive filter algorithm.

According to this principle, a new variable step LMS adaptive algorithm is proposed here. The new algorithm adjusts the step size according as the convergence ratio of MSE, and analyses the origin of the mutation according to the correlation between reference signal and output error which following relevant strategy of modifying the step size.

2 The New Variable Step LMS Algorithm Based on the Convergence Ratio of MSE

As Fig.3 showed, for general LMS adaptive filter algorithm, the MSE convergence ratio fluctuates according to a positive value during initial convergence; when the MSE is nearly in the steady-state, the fluctuation center of convergence ratio will fall to zero in a fast speed; besides, the fluctuation center is proportion to the step coefficient μ.According to these feature of the MSE convergence ratio, the new variable step LMS adaptive algorithm proposed in this paper recursion is given by

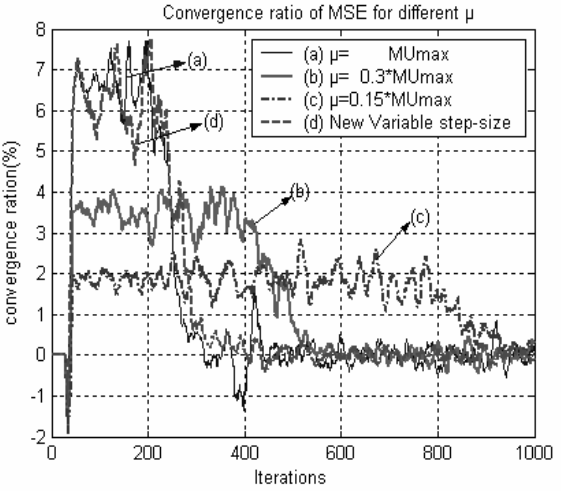


Fig. 3. The convergence ratio of MSE for different μ

$$e(n) = d(n)- X^T(n)W(n) \tag{4}$$

$$\mu(n+1)=\begin{cases} \mu(n) & \Delta e^2(n)>0 \\ \mu(n)\times\alpha & \beta<\Delta e^2(n)<0 \\ \mu(n)\times\alpha & \beta<\Delta e^2(n)<0 \text{ and } MRxe(n)<50\times ERxe(n) \\ \lambda max & \beta<\Delta e^2(n)<0 \text{ and } MRxe(n)>50\times ERxe(n) \end{cases} \tag{5}$$

$$W(n+1) = W(n) + 2\mu(n+1)e(n)x(n) \tag{6}$$

Where

$$\Delta e^2(n)=[e^2(n-1)-e^2(n)] / e^2(n-1)\times 100 \tag{7}$$

is the the MSE convergence ratio, α is the attenuation ratio of the step coefficient μ, β is the boundary point of $\Delta e^2(n)$ mutation; $R_{xe(n)}(m)$ is the cross-correlation function between the $X(n)$ and $e(n)$,and $MRxe(n)=\max(R_{xe(n)}(m))$,$ERxe(n)=\text{mean}(R_{xe(n)}(m))$; L is the length of the order of adaptive filter.

$$R_{xe(n)}(m)=\sum_{p=n-L+1}^{n} x(p)e(p+m) \tag{8}$$

3 Algorithm Simulation and Performance Analysis

Now we will analysis the performance of the new algorithm according to the simulation results, and the simulation conditions as following:

(1) Input signal $X(n)$ is Gaussian noise, interference $V(n)$ is also Gaussian noise but uncorrelated with $X(n)$; (2) FIR coefficients of unknown system is a random row matrix with 5 elements. At the time of sampling point 600,unknown system is time-varying and the FIR coefficients is also a 1×5 random row matrix;(3)Order of adaptive filter L=10;(4) The step-size factor in Equation(5):α=0.9,β=-50;(6) The sample points are 1000 and results are obtained by averaging over 500 independent runs.

1) The convergent speed and stable state maladjustment

The curve (d) in Fig.2 is the error convergence curve of improved algorithm. Compared with the curve (a)(b)(c),it shows that this algorithm has faster convergence speed and lower maladjustment of steady-state than the general LMS algorithm. It can solve the problem of the convergence speed conflicts with the accuracy of steady-state.

2) The trace ability of time-varying system and resistance to input disturbance

In this part, the trace ability of time-varying system and resistance to input disturbance of different adaptive filter algorithm is analyzed, which is showed in Fig.4 and Fig.5 .

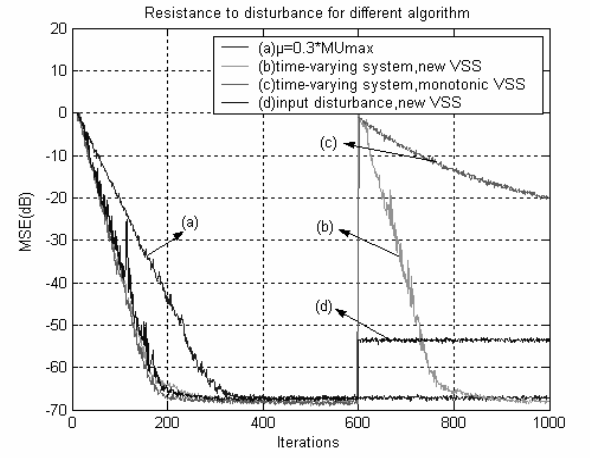


Fig. 4. The convergent speed of different algorithm

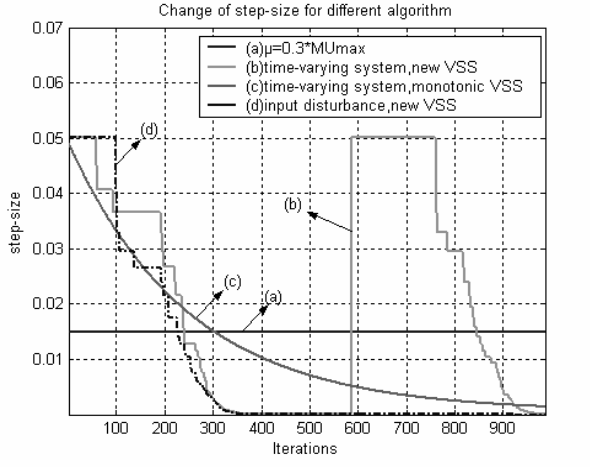


Fig. 5. The change of step-size for different algorithm

3) Calculation analyse

Noticeably, the contradiction between the convergent speed and stable state maladjustment is banished, and the robusticity increases, but the increase of the operation is very small. The new VSS algorithm increases one difference operation and one division operation comparing to general LMS algorithm, which can be found easily from Equation(4~8).Although the cross-correlation function between $X(n)$ and $e(n)$ is also used in this new algorithm, it will happen while the mutation of the MSE arises. The selective operation of correlation function is a big improvement compared to the VS-NLMS proposed in reference [7].

4 Conclusion

The simulation results prove that the new VSS LMS algorithm proposed in this paper is effective. It is only small increase of operation to general LMS algorithm, which speeds up the convergence, reduces the steady state maladjustment and improves the trace ability of time-varying system and resistance to input disturbance.

References

1. Dimitris, G.M., Vinay, K.I., Stephen, M.K.: Statistical and Adaptive Signal Processing, pp. 499–501. Publishing House of Electronics Industry, Beijing (2003)
2. Widrow, B.: E Walach Adaptive Inverse Control. Xi'an Jiaotong University Publishing House, Xi'an (2000)
3. Sock, D.T.M.: On the Convergence Behavior of the LMS and the Normalized LMS Algorithms. IEEE trans. on signal processing 41(9), 2811–2825 (1993)
4. Tang, J., Zhang, J.S.: An Improved Variable Step Size LMS Adaptive Filtering Algorithm and its Analysis. In: The ICCT 2006, vol. 11, pp. 27–30 (2006)
5. Kwong, R.H., Johnston, E.W.: A Variable Step Size LMS Algorithm. IEEE Trans on Signal Processing 40, 1633–1642 (1992)
6. Zhang, Y.G., Li, N., Jonathon, A.C.: A New Gradient Based Variable Step-Size LMS Algorithm. In: 8th International Conference on Signal Processing, pp. 16–20 (2006)
7. Gan, W.S.: Designing A Fuzzy Step Size LMS Algorithm. IEE Image Signal Process 144(5), 261–266 (1997)
8. Li, S.T., Wan, H.: A New Variable-Step-Size LMS Algorithm and Its Application in FM Broadcast-Based Passive Radar Multi-Path Interference Cancellation. In: 2007 Second IEEE Conference on Industrial Electronics and Applications, vol. 2, pp. 2124–2128 (2007)
9. Li, Y.: A Modified VS LMS Algorithm. In: 9th International Conference on Advanced Communication Technology, pp. 615–618 (2007)
10. Wang, Y., Zhang, C., Wang, Z.H.: A New Variable Step Size LMS Algorithm with Application to Active Noise Control. In: IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings, vol. 5, pp. 573–575 (2003)

Saliency-Based Image Quality Assessment Criterion*

Qi Ma and Liming Zhang

Department of Electronic Engineering, Fudan UniversityNo.220,
Handan Road, Shanghai, 200433, China
ma.lance@gmail.com, lmzhang@fudan.edu.cn

Abstract. Image quality assessment (IQA) is a critical issue in image processing applications, but traditional criteria based on the differences between reference and distorted images do not correlate well with perceived quality. MSSIM and VIF criteria proposed recently are regarded as excellent models in comparison with others, but they only consider local features and ignore some global concepts. In this paper, we propose an improved method that adds global saliency features to the criteria of MSSIM and VIF. For the sake of reducing the computational complexity, we propose a simpler and faster method to extract the saliency map. Experimental results for a set of intuitive examples as well as validation on a database of 779 images with different distortion types demonstrate that the improved IQA criteria can get a better performance than their original forms.

Keywords: Image quality assessment, Saliency map, Human visual system, Mean structural similarity, Visual information fidelity.

1 Introduction

Recently, image quality assessment (IQA) is becoming more attractive in most of image processing applications since traditional full-reference IQA, such as peak signal-to-noise ratio (PSNR) or mean-square error (MSE) does not accord with human judgment well in some cases. The basic idea to assess perceptual quality of images is to score these distorted images by human observers. However, such subjective evaluation is usually cumbersome and expensive. In fact, it is not practical in image processing systems. Therefore, the motivation for IQA research is to get a simple objective evaluation of image quality which is consistent with subjective human judgment.

There has been a good deal of literature proposing some IQA methods and criteria. C. Bovic etc. proposed two kinds of criteria in 2004-2006[1][2][3] that assume the perfect or reference images are available. One criterion is based on structural similarity [1]. They considered that natural images are highly structured and proposed a criterion to capture the structure's differences between the reference and distorted images. This method divides the reference and test images into many overlapped blocks, and then the comparison of each block pair is based on three metrics: luminance, contrast and structural correlation. Final criterion combines the three metrics

* This work is supported by the National Nature Science Foundation of China (60571052) and Shanghai Leading Academic Discipline Project (B112).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1124–1133, 2008.
© Springer-Verlag Berlin Heidelberg 2008

on the average of all the block pairs, which is called Mean structural similarity (MSSIM). Another criterion called visual information fidelity (VIF)[3] is based on mutual information between input and output of human vision system (HVS) channel for both reference image and distorted image. VIF has two versions for spatial domain and wavelet domain. It is also based on blocks for different scales or sub-bands. An image database with five-types of distorted images and original reference images is provided to test the two criteria. All the distorted images are ranked difference mean opinion score (DMOS) by human observers. The results show that MSSIM and VIF have higher performance than other methods such as PSNR, JND-metrix etc.

Attention selection is a very important property in human visual system. Eyes of human usually attend to edges with high contract or salient areas that are different from their circumambience. Saliency detection is a key stage to object detection, which also should be considered in IQA. Unfortunately, existing computational models to simulate visual attention, such as neuromorphic vision C++ toolkit (NVT)[4] and saliency toolbox(STB)[5], are too complicated and rely on the choice of parameters. So they are very difficult to be implemented in IQA. This paper proposes a simple and fast method to extract saliency map in reference image, then uses the saliency map as weights for MSSIM or VIF. Since MSSIM and VIF are both based on local features, they lose the global information of high importance. Attention saliency just grasps possible object information which would improve the performance of MSSIM and VIF. This idea also can be used for any other IQA in spatial domain.

This paper is organized as follows: in section 2 we give an example to explain the power of Attention selection for IQA, section 3 introduces our proposed model to extract attention saliency map, section 4 presents our saliency-based IQA, simulation results and a conclusion are given in section 5 and 6, respectively.

2 Attention Areas and IQA

IQA tries to develop an objective image quality metric that approaches human judgment. Since attention is a very popular principle in human vision, we should consider it in IQA problem. Given an image, human vision would pay more attention to these interesting regions and less attention to other regions of the image. Obviously, the two types regions should be treated differently in IQA. Saliency map is a scalar map that indicates the human attention areas of the original image. The locations with high value in the saliency map represent the regions which can attract more attention. Note that saliency map is different from the edges extraction, since only novel information will pop out on a saliency map. Repetitive texture or objects with obvious edges will be inhibited. Saliency map represents the global information of the image.

Let us show an example in Fig.1 named “Rapids”. Fig.1.a is the reference image, Fig.1.b is its saliency map, and Fig.1.c, Fig.1.d, Fig.1.e and Fig.1.f are distorted images polluted by blur or noise. The difference is that Fig.1.d and Fig.1.f are polluted in less salient areas (river-blurred and river-noised), and Fig.1.c and Fig.1.e are polluted in salient areas (boat blurred and boat-noised). We can easily distinguish the quality of two images pairs — the quality of Fig.1.d is better than Fig.1.c. Fig.1.e is better than Fig.1.f. However, MSSIM and VIF (in spatial domain) get the similar assessment index for Fig.1.d and Fig.1.c (note that higher index denotes higher

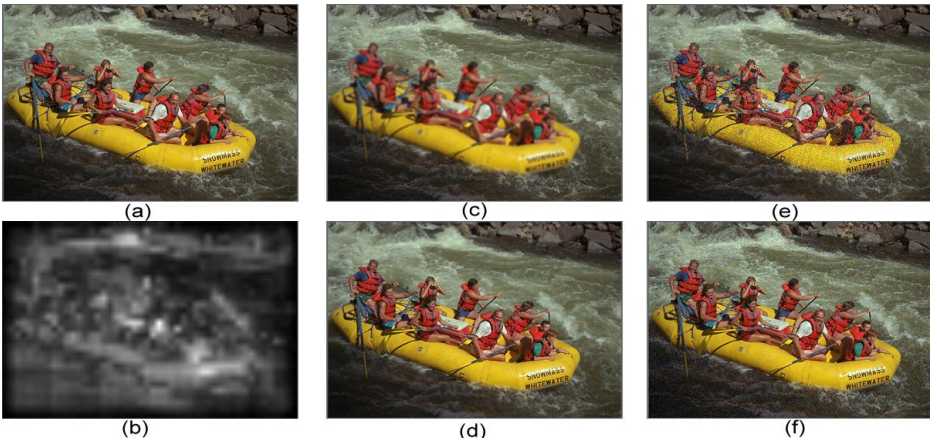


Fig. 1. Comparison of images with different perceptual qualities. (a) Original "Rapids" image. (b) Saliency map of (a) by STB. (c) Image with blurred pollution in attention area, MSSIM=0.85, VIF=0.63. (d) Image with blurred pollution in less attention area, MSSIM=0.84, VIF=0.63. (e) Image with noised pollution in attention area, MSSIM=0.80,VIF=0.63. (f) Image with noised pollution in attention area, MSSIM=0.69, VIF=0.61.

quality seen in section 4), Fig.1.e and Fig.1.f obtain the result contrary to human evaluation (the MSSIM of Fig.1.e is 0.80 while the MSSIM of Fig.1.d is 0.69). This shows that even the two most outstanding IQA criteria fail in this case just because they ignore human's visual attention mechanism.

How to get saliency map from original image? Models like NVT and STB are too complex to be implemented. In section 3 we give a simple method to extract saliency map from a reference image.

3 Saliency Map Based on Phase Spectrum

Owing to the deficiencies of existing vision attention algorithms, we consider a novel visual attention model from another point of view. What can attract human attention from a given image? First, discontinuous parts such as objects' edges or strong changes in grey value, which includes a lot of sinusoidal frequency component, will be considered as candidates of attention; Second, novelty information can be distinguished. Our work, that will be published in CVPR2008[6], indicates that phase spectrum of 2D-Fourier transform results in the saliency map because phase spectrum specifies where each of the sinusoidal components resides within the image. These locations for novelty information or high contrast just include many sinusoidal components. We propose a model called phase spectrum of Fourier transform (PFT) algorithm. Its steps are shown as follows: Given an image $I(x, y)$

step1 $$f(u,v) = F(I(x,y)) \tag{1}$$

step2 $$P(u,v) = P[f(u,v)] \tag{2}$$

step3 $$S(x,y) = g(x,y) * F^{-1}[\exp(i \bullet P(u,v)] \quad (3)$$

Where F and F^{-1} denote the Fourier Transform and Inverse Fourier Transform respectively. $P(f)$ denotes the phase spectrum of the image. For better visual effect, we smoothed the saliency map with a Gaussian filter $g(x)$ ($\sigma = 8$). This algorithm for extracting the saliency map only cost .0.01-0.02s for one image. Fig.2 gives a comparison between PFT and STB (that cost 4-5s for one image) which shows that PFT achieves a similar result to STB, while saves considerable computing time owing to the low computation complexity.

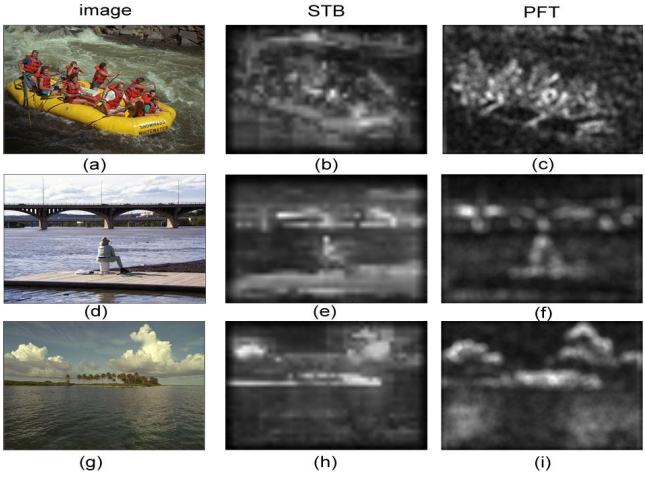


Fig. 2. Comparison between STB and PFT.(a)(d)(g) Reference image, (b)(e)(h) The saliency maps of (a)(d)(g) by STB, (c),(f),(i) The saliency maps of (a)(d)(g) by PFT.

4 Saliency-Based IQA Measures

How to introduce image saliency information into current IQA measures is a remarkable question. For reducing computational complexity, in this paper we treat the saliency map as weights to adjust the contribution of every small block's (or even pixel's) index in the final IQA metric.

4.1 Saliency-Based MSSIM (SMSSIM)

The SSIM index proposed by Wang etc. is to calculate the local structural information in an image as well as the average luminance and contrast between reference and distorted images. SSIM first divides reference and distorted images into M overlapped blocks. For each block we can calculate the corresponding SSIM index. Let X be a block in reference image, Y be the relevant (same location with X) block in the distorted image, then the SSIM index of this block is

$$SSIM(\mathrm{x},\mathrm{y}) = l(\mathrm{x},\mathrm{y})\cdot c(\mathrm{x},\mathrm{y})\cdot s(\mathrm{x},\mathrm{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2+\mu_y^2+C_1}\cdot\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2+\sigma_y^2+C_2}\cdot\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \tag{4}$$

Where $l(\mathrm{x},\mathrm{y})$, $c(\mathrm{x},\mathrm{y})$ and $s(\mathrm{x},\mathrm{y})$ represent local image luminance, contrast, and structural correlation respectively, μ_x, μ_y are means of grey value in block X and Y respectively, σ_x^2, σ_x^2 are variances of X and Y, and σ_{xy} is the cross-covariance between X and Y, the constants C_1, C_2, C_3 are use to stabilize the algorithm. It is obvious that if the distorted and referenced block images are the same, Eq.(4) approaches to one. In a similar way, we divide the saliency map *SMap*, obtained from reference image by PFT algorithm, into the overlapped blocks of the same size as X or Y. Let S represent the block of *SMap*, at the same location with X and Y, then saliency-based SSIM can be defined as

$$SSSIM(\mathrm{x},\mathrm{y},\mathrm{s}) = SSIM\cdot\mu_s \tag{5}$$

Where μ_s is the normalized mean of grey value in block S. Since the mean values of *SMap* blocks in different locations are different, the μ_s adjusts the weight of each SSIM index. By averaging all SSSIM indexes, we obtain saliency-based MSSIM,

$$SMSSIM = \frac{1}{M}\sum_{j=1}^{M} SSSIM(\mathrm{x}_j,\mathrm{y}_j,\mathrm{s}_j) \tag{6}$$

Where is the total number of blocks. Since μ_s is normalized, Eq.(6) still approaches to one for no distorted image.

4.2 Saliency-Based VIF (SVIF)

The VIF proposed by Hamid R. Sheikh etc. is another recent IQA criterion that consistently outperforms almost all other criteria. It treats the IQA as an information fidelity problem. VIF has two types: wavelet domain version and pixel domain version. Considering that wavelet domain version VIF is more complex and our saliency map is based on pixel domain, in this paper we only discuss the pixel domain version. Based on the information fidelity idea, the researchers take natural image as an information source and the reference image received from HVS channel (an additive white noise channel (AWGN)) as information source plus noise.

If c is a block vector from a given location in the natural image at a certain scale level, e is the block c perceived by human eyes as reference block image, which can be represented as $e = c + n$, where n is additive noise. The distorted block image perceived by the eyes is $f = d + n = gc + v + n$, where $d = gc + v$ is the block vector from the relevant distorted image, g represents blur distortion and v is noisy distortion by additive white Gaussian noise. Let $blockMI(c,e) = \mathrm{I}(c,e)$ represent the mutual information between c and e of one block. By summing all block's mutual information, we

have $ScaleMI = \sum_{i=1}^{blocks} \mathrm{I}(c_i; e_i)$, the mutual information in one scale of image, where i denotes all possible blocks in the scale. By summing all scale level's mutual information, we get $MI(nature, ref) = \sum_{j=1}^{scales} \sum_{i=1}^{blocks} \mathrm{I}(c_{i,j}, e_{i,j})$, the total mutual information between natural image and perceived reference image which represents how much information of natural image can be extracted from the reference image. For the same reason, we get $MI(nature, dis) = \sum_{j=1}^{scales} \sum_{i=1}^{blocks} \mathrm{I}(c_{i,j}, f_{i,j})$, the total mutual information between natural image and perceived distorted image which denotes how much information of natural image that the brain extracts from the distorted image. VIF is the ratio between distorted image information and reference image information.

$$VIF = MI(nature, dis) / MI(nature, ref) \tag{7}$$

It is obviously VIF in pixel domain is also a metric based on small blocks at different scales. At each scale level, we obtain the corresponding saliency map $SMap_j$ (j denotes the scale level), and then divide $SMap_j$ into the same size blocks as MI. Let $\mu_{i,j,s}$ represent a normalized mean of the i^{th} block in the j^{th} scale of $SMap$ with same size and location as natural sub-image c, reference sub-image e and distorted sub-image f. Saliency-based pixel domain VIF can be defined as

$$SVIF = \frac{WeightedMI(nature, dis)}{WeightedMI(nature, ref)} = \frac{\sum_{j=1}^{scales} \sum_{i=1}^{blocks} (blockMI(c_{i,j}, e_{i,j}) \bullet \mu_{i,j,s})}{\sum_{j=1}^{scales} \sum_{i=1}^{blocks} (blockMI(c_{i,j}, f_{i,j}) \bullet \mu_{i,j,s})} \tag{8}$$

4.3 Other Saliency-Based IQA Method

Our method can be easily extended for any other IQA in spatial domain. We take mean squared error (MSE) and peak signal-to-noise ratio, the most widely used IQA indexes, for example. Given two images, reference image $X = \{x_i \,|\, i = 1, ..., N\}$, and distorted image $Y = \{y_i \,|\, i = 1, ..., N\}$. Let $SMap_i$ be the value of the ith pixel on the saliency map $SMap$, we propose the saliency-based MSE metric as

$$SMSE(X, Y, SMap) = \frac{1}{N} \sum_{i=1}^{N} \left((x_i - y_i)^2 \cdot SMap_i \right) \tag{9}$$

Considering MSE is based on pixel level rather than blocks, saliency map can easily play role in this case. The values of the saliency map adjust the weight directly. Then saliency-based PSNR is

$$SPSNR(X,Y,SMap)=10\log_{10}\left(\frac{L^2}{SMSE(x,y)}\right) \tag{10}$$

Where L is the image dynamic range (typically, L=255).

5 Experimental Results

In this section, we test our algorithm on an intuitive example and a database from [7].

5.1 Intuitive Example

Table1 shows the quality assessment results of distorted images in Fig.1 using PSNR, SPSNR, MSSIM, SMSSIM, VIF and SVIF indexes (where px denotes VIF in pixel domain). From Fig.1, we know that the qualities of Fig.1.d and Fig.1.f are better than that of Fig.1.c and Fig.1.e in human judgment. The saliency-based IQA correspond with human perceptual quality well, but their original IQA criteria fail in this case.

Table 1. Results for images of Fig.1 using different quality indexes

Image	PSNR	SPSNR	MSSIM	SMSSIM	VIF(px)	SVIF(px)
Fig.1.c	28.54	23.77	0.85	0.80	0.63	0.51
Fig.1.d	28.61	31.89	0.84	0.89	0.63	0.69
Fig.1.e	23.67	21.06	0.80	0.72	0.63	0.48
Fig.1.f	23.82	26.85	0.69	0.80	0.61	0.65

5.2 Validation on Database

In order to validate our idea, we test our proposed IQA criteria on a database from [7]. Twenty-nine high-resolution color images were distorted using five distortion types: JPEG2000, JPEG, white noise in the RGB components, Gaussian blur, and transmission errors in the JPEG2000 bit stream using a fast-fading Rayleigh (FF) channel model. Each distorted image was rated by human observers who provided their rank of quality on a continuous linear scale. The raw scores were converted to difference scores and then to Z-scores, scaled back to 1–100 range, finally to a difference mean opinion score (DMOS). So the database contains a total of 982 images (779 distorted images) with difference mean opinion score (DMOS) for each distorted image.

5.3 Simulation Details

All IQA criteria operated upon the luminance component only. Considering that the saliency maps achieved by different images may not have the same intensity level which will severely affect the final result, we first uniform the saliency map as follows

$$SMap(x_i) \leftarrow SMap(x_i)/\left(\frac{1}{N}\sum_{i=1}^{N} SMap(x_i)\right) \tag{11}$$

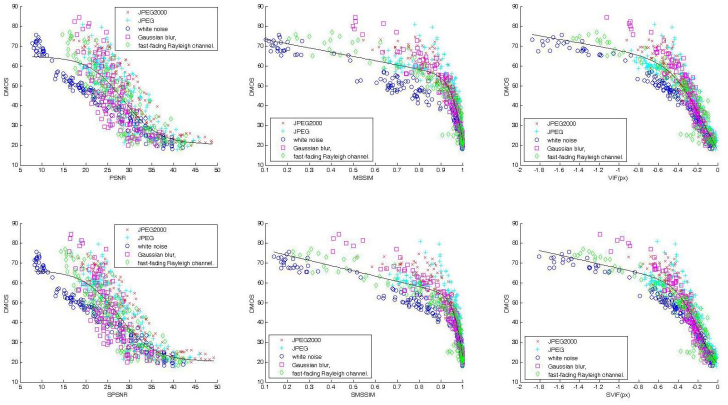


Fig. 3. Scatter plots for the six objective quality criteria: PSNR, SPSNR, MSSIM, SMSSIM, log(VIF) and log(SVIF) (VIF and SVIF are in pixel domain. The distortion types are: (×) JPEG2000, (+) JPEG, (°) white noise, (box) Gaussian blur, and (diamond) transmission errors in JPEG2000 stream over fast-fading Rayleigh channel.

Where $SMap(x_i)$ is the gray value of pixel x_i, N denotes the total pixel number in saliency map. This uniform process guarantees the mean of gray value of all pixels in saliency map equal to 1 for SPSNR, and the mean of all μ_s equal to 1 approximately for SMSSIM and SVIF(px).

Three IQA and relevant saliency-based IQA methods are applied to all the distorted images in the database (779 images). We can get the index of MSSIM, VIF (px), PSNR, saliency-based index SMSSIM, SVIF and SPSNR. Fig.3 shows the six scatter-point plots of DMOS vs. six IQA indexes for five-type distortions. A regressive curve in each scatter plot denotes the relation between objective and subjective IQA. The standard deviation between scatter points and regressive curve can denote the performance of IQA methods. Small standard deviation denotes the better performance of IQA. It is obvious from Fig.3 that PSNR is the worst and SVIF is the best among the six IQA methods. All saliency-based IQA are better than their original ones. From Fig.3, we can see that each regressive curve gives a relation between subjective assessment (DMOS) and that objective IQA. For given an IQA index, we can predict DMOS from the corresponding regressive curve. In order to quantitative comparison, we transfer Fig.3 to Fig.4. In Fig.4 y-axes is the DMOS scores by human and x-axes is the predicted DMOS from regressive curve. There are five quantitative indexes from Fig.4 to compare different IQA methods. Correlation coefficient (CC) provides an evaluation of prediction accuracy in Fig.4. If all predicted DMOS (Quality) in database equals to human DMOS, CC=1. In general, CC<1. MAE (mean absolute prediction error), RMS (root mean square prediction error) are error between objective/subjective scores, OR (outlier ratio the percentage of the number of predictions outside the range of ±2 times of the standard deviations which is a measure of prediction consistency) and SROCC (Spearman rank-order correlation coefficient which is a measure of prediction monotonicity) between objective and subjective scores.

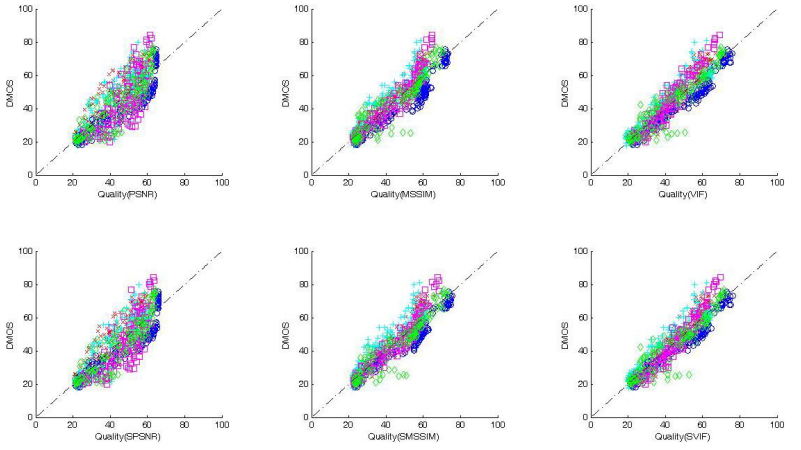


Fig. 4. Scatter plots for the quality predictions by the six methods after non-linear regression for quality calibration: PSNR, SPSNR, MSSIM, SMSSIM, log(VIF) and log(SVIF). The distortion types are: (×) JPEG2000, (+) JPEG, (°) white noise, (box) Gaussian blur, and (diamond) transmission errors in JPEG2000 stream over fast-fading Rayleigh channel.

5.4 Testing Results

Table 2 shows the testing results of five quantitative indexes for all 779 images belonging to five distortion types in the database. It is same as scatter plots in Fig.3 and Fig.4. SVIF has the highest CC and SROCC and the smallest MAE, RMS and OR, so SVIF accords with human's opinions better. PSNR is the worst one. It also has been seen that SPSNR, SMSSIM and SVIF outperform their original ones. It is reasonable that they considered Human attention in image quality assessment.

Table 2. Validation scores for different quality assessment methods. The methods tested were PSNR, SPSNR, MSSIM, SMSSIM, VIF(px) and SVIF(px).The methods were tested against DMOS from the subjective study after a nonlinear mapping.

MODEL	CC	MAE	RMS	OR	SROCC
PSNR	0.824	7.324	9.123	0.118	0.819
SPSNR	0.854	6.871	8.640	0.093	0.841
MSSIM	0.909	4.986	6.491	0.036	0.909
SMSSIM	0.925	4.627	6.166	0.026	0.923
VIF(px)	0.923	4.786	6.130	0.027	0.925
SVIF(px)	0.934	4.501	5.886	0.022	0.933

6 Conclusion

In this paper we present a saliency-based strategy into image quality assessment, and propose three saliency-based approaches by modifying the metrics of three most

widely used IQA methods. These new approaches give better performances than their original forms both in an intuitive example and on a database of 778 images with different distortion types. A simple Fourier phase spectrum can easily get saliency map. So the improved IQA method can be automatically implemented easily. In the light of remarkable performance of saliency-based IQA, it is worthy of adding saliency detection strategy to IQA to improve their performance.

However, in this paper we only treat saliency map as some kind of weight on original IQA formulation. How to add attention area to other domain, such as VIF in wavelet domain, still remains space for our further work.

References

1. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image Quality Assessment: from Error Visibility to Structural Similarity. IEEE Transactions on Image Processing 13(4), 600–612 (2004)
2. Sheikh, H.R., Bovik, A.C., de Veciana, G.: An Information Fidelity Criterion for Image Quality Assessment Using Natural Scene Statistics. IEEE Transactions on Image Processing 14(12), 2117–2128 (2005)
3. Sheikh, H.R., Bovik, A.C.: Image Information and Visual Quality. IEEE Transactions on Image Processing 15(2), 430–444 (2006)
4. Itti, L., Koch, C., Niebur, E.: A Model of Saliency-based Visual Attention for Rapid Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(11), 1254–1259 (1998)
5. Walther, D., Koch, C.: Modeling Attention to Salient Proto-objects. Neural Networks 19, 1395–1407 (2006)
6. Guo, C.L., Ma, Q., Zhang, L.M.: Spatio-temporal Saliency Detection Using Phase Spectrum of Quaternion Fourier Transform. In: IEEE Conference on Computer Vision and Pattern Recognition, p. 116 (2008)
7. http://live.ece.utexas.edu/research/quality

Crowd Segmentation from a Static Camera

Bin Lai , Deng Yi Zhang *, Zhi Yong Yuan , and Jian Hui Zhao

Department of Computer Science,
Wuhan University, Wuhan, Hubei 430079, China
Bin.y.lai@gmail.com, dyzhangwhu@163.com

Abstract. This paper presents a more enhanced and efficient method for crowd segmentation from background subtracted images using active basis model associated detection cascade. Firstly, the problem is significant because the case of inter-human occlusion usually appears in the image which may disturb the result of tracking and recognition, consequently debases the effect of the whole surveillance system. Secondly, the problem is challenging because the state space formed by the number, positions, and articulations of people is large, and the noisy may confluence the achievement of the shape of the crowd of background subtracted. We combine some effective methods into one system for improving the hit rate in crowded scene, such as the head position estimation, active basis model, etc. Especially, using the active basis model to verify the result from detection cascade, the system gives the excellent performance for detecting human in crowded scene.

Keywords: Crowd Segmentation, Active Basis Model, Detection Cascade.

1 Introduction

Detecting and segmentation of the pedestrians in a scene are important for some applications areas, such as security, safety, smart interfaces, etc. These applications always adopt fixed cameras. In the situation of fixed cameras, background subtraction is a important method to extract the moving pixels (foreground). If objects are sparse in the scene, each connected component of the foreground (blob) usually corresponds to an object, though the blobs may be fragmented due to low contrast or partial occlusion by scene objects or may contain non-human pixels caused by shadow.

When the density of the objects increases, it is common that several objects form one big blob. Therefore the blobs do not directly provide the object level description that is needed for human event inference. Our goal in this paper is to segment the foreground into individual human objects which may overlap with each other.

Some work has been done to segment or track multiple overlapping humans. In [1], peaks in the vertical histogram of the blob are used to help locate the positions of the heads. In [2], vertical peaks on the foreground boundary are used to locate the positions of the heads. In [3], humans are assumed to be isolated as they enter the scene so that a human specific color model can be initialized for segmentation when occlusion occurs. [2] also use the initialized human specific model to help in tracking through occlusion. Leibe et al [9] describes an interest-point and

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1134–1140, 2008.
© Springer-Verlag Berlin Heidelberg 2008

local feature descriptor based detector, followed by global grouping constraints to detect humans. Wu and Nevatia [7] describe a parts based human detector and extend it to handle multiple humans. Such approaches are complex, computation intensive, and it is not clear how well they can generalize to arbitrary surveillance situations and people appearances. Face detection may not be effective when the humans don't face the camera or when the image size of the human is small (e.g., [5]). Detecting human directly (e.g., [8]) has been limited to restricted viewpoints (frontal or back). Zhao and Nevatia [4] describe a generative process where the parameters include the number of people, their location and their shape. They use Markov Chain Monte Carlo (MCMC) to achieve global optimization by searching for maximum likelihood estimates for the model parameters. The main issue with these approaches is the complicated and high dimensional parameter space and the subsequent slow process of searching for the best parameters. Lan Dong and Vasu [6] choose Fourier descriptors for representing a shape and built an indexing function that mapped observed descriptors to candidate parameter sets explaining the shape.

2 Our Method

Through combining several effective human detection methods into a system, we improve the detection rate and depress the false alarm rate. Our system is comprised of four components which collaborate to detect the moving group and segment it to individuals. The four components are (Fig.1 shows the architecture of the system):

1) Foreground Region Detector: to detect these foreground regions containing moving objects in each frame.
2) Head Position Detector: to estimate the position of heads in the foreground regions.
3) Upper Body Detector: to detect the upper body in the foreground region by using a haar-like feature based detection cascade.
4) Head-shoulder Detector: to verify the detected head-shoulder position by using the Upper Body Detector.

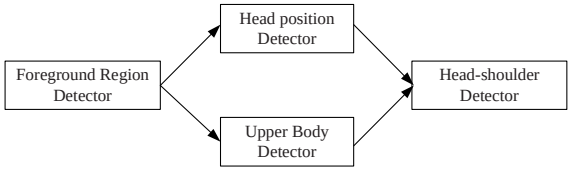


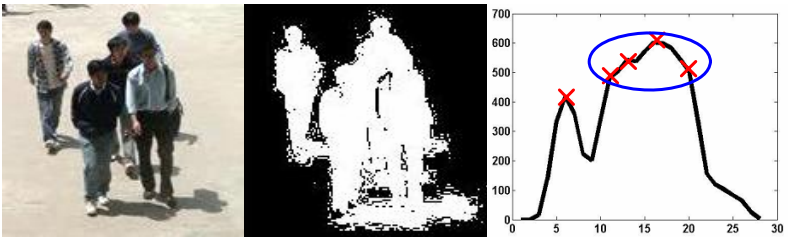
Fig. 1. The architecture of the detection system includes the four parts

2.1 Foreground Region Detector

Background estimation and change detection are quietly developed areas in visual surveillance [11]. There are many sophisticated algorithms which can compensate for various extraneous effects such as flickering, lighting changes, weather etc., [3] surveys the

development of the image change detection algorithm. But the result of the change detection process may be not good in some situation, such as strong reflections, shadows and regions where foreground and background get the similar colors, especially, in front of the infection of strong shadows and lighting in the indoor cases, the detection of foreground region always fails, so Foreground Region Detector just can be adopt in the outdoor cases. In this paper, we just discuss the outdoor cases.

For increasing the rate of detection and deducing the computational cost, at first, we should make background subtraction, then the algorithm just focuses on the region of the foreground which is consisted of the moving objects, assuming that humans are the only the moving objects in the image. Then it uses a low-pass filter to compensate for noise in the motion image. Fig.2(b) shows the binary image after background subtraction.



(a) (b) (c)

Fig. 2. The silhouette based features. (a) input image, (b) detect foreground region, (c) Vertical projection, the red X shows all the most possible positions of heads, and the blue ellipse shows the region of heads.

2.2 Head Position Detector

We have implemented the algorithm to detect head positions in the interior of the foreground blobs which is mentioned in the W4 system by Haritaoglu et al [1]. The algorithm creates a vertical pixel histogram of these foreground regions in the binary image acquired by foreground region detector, then scans the vertical histogram for peaks, using camera calibration to analyze whether these might be heads of people (Figure.2.(c)). The algorithm records all the possible positions of the head to later use by the head-shoulder detection. Its implementation in the Head Position Detection is obviously simple, fast and effective which can be competent for fast detection.

2.3 Head-Shoulder Detector

The head-shoulder shape is important feature which is used to detect the occluding human in crowded scene. The reliable method to detect the head-shoulder feature is active basis model [12], which fuses the shape models such as active contours [14] and active appearance model [13]. We can learn the active basis model by the shared sketch algorithm on a training set of head-shoulder patches. According to the calibration of the camera, we can adopt the appropriate scale of the learned model, and the head position detector can associate the active basis model to locate the initial position,

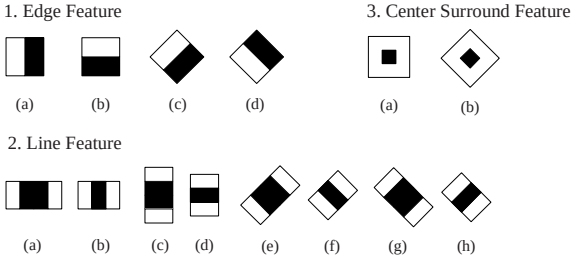


Fig. 3. Feature prototypes of simple haar-like and center-surround features

finally we will get the log of likelihood rate at all the possible positions using the template match. Through choosing an appropriate threshold, the algorithm can discriminate whether there exists a pedestrian. The drawback of the method is huge computation which constrains the real-time applications.

2.4 Upper Body Detector

For improving the hit rate of the whole system, besides the module of the Head-Shoulder Detector, the module of the Upper Body Detector supports another path to detect the human in the crowded scenes.

We adopt a cascade of boosted classifiers for human detection. The feature pool includes the over-complete set of haar-like features used in [10], and their fast computation scheme is proposed in [15]. All the feature prototypes are shown in Fig.3:

We use gentle adaboost as our basic classifier. Adaboost is a powerful learning algorithm which combines many weak classifiers to a strong classifier by re-weighting the training samples. Our set of weak classifiers is all classifiers which use one feature from our feature pool in combination with a simple binary thresholding decision. With increasing stage number, the number of weak classifiers increases, which are needed to achieve the desired false alarm rate at the given hit rate.

A cascade of classifiers is a degenerated decision tree where at each stage a classifier is trained to detect the pedestrian. In our case, each stage was trained to eliminated 50% of the non-upper body patterns while falsely eliminating only 0.5% of the upper body patterns; 22 stages were trained.

As the features can be computed at any position and any scale, the searching region is restricted to the foreground region.

3 Experiments and Evaluation

3.1 Learning Active Basis Model

Parameter values. The number of the basis elements is 50, the Size of Gabor wavelets = 17 × 17. The orientation α takes K = 15 equally spaced angles in $[0, \pi]$, the orthogonality tolerance is ε = 0.1. The threshold is T = 16 in the threshold transformation. The

Fig. 4. Active basis formed by 50 Gabor wavelet elements. The first plot displays the 50 elements, for each of the other 17 pairs, the left plot is observed image, and the right plot displays the 50 elements resulting from locally shifting the 50 elements in the first plot to fit the corresponding observed image

shift along the normal direction $d_{m,i} \in [-6,6]$. The shift of orientation $\delta_{m,i} \in [-1,1]$ angles out of K = 15 angles.

Given a training set of M = 93 head-shoulder images, we can learn the active basis model (Figure.4 shows this model) using the shared sketch algorithm.

3.2 Detecting by Using Active Basis Model

Active basis model can detect the pedestrians in crowd effectively (Figure.5 shows the process). For an input image, we can apply the algorithm at multiple resolutions of the input image. Then we can choose the resolution that achieves the maximum score as the optimal resolution, finally, we sketch the image at this resolution.

The disadvantage of the novo model is large computation, we have to calculate the Gabor filter on the image for each orientation, and the several different scales of sliding window will result in increasing the computation, so the method is not adopted in real-time surveillance at present. With the speed of hardware increasing, the ability of this method will appear for object detection.

(a) (b) (c) (d) (e)

Fig. 5. Human detecting by using active basis

3.3 Performance of the System

At first, we use the haar-like feature based detection cascade to detect the upper bodies, we can abide the higher false alarm rate for the higher hit rate at this step, as the active basis model has the excellent validity check, we can use it to debase the false alarm rate. We compare the the haar-like feature based detection cascade with the integrated system include the two models, Figure.6 shows the ROC for comparison. While the false positives per window is 1/10000, the hit rate of our system is 87.3%, which is obviously higher than 82.4% of the detection cascade.

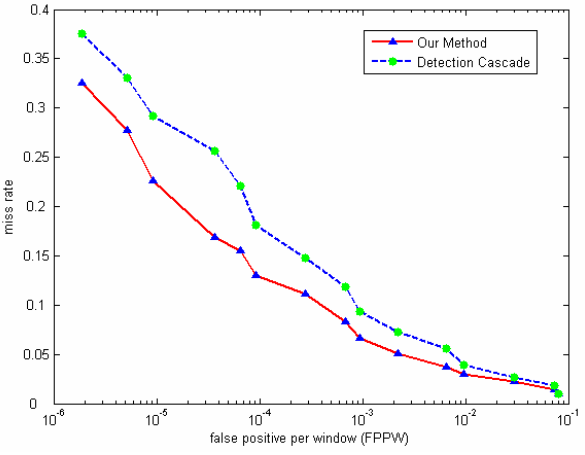


Fig. 6. The ROC describes that the active basis model improve the performance of human detection obviously

4 Conclusion

We present a more enhanced and efficient method for crowd segmentation from background subtracted images by using active basis model associated detection cascade. As the head-shoulder feature is very useful for the crowd segmentation, and the traditional methods, such as snake algorithm, are easily disturbed by unobvious edge. The active basis model is more reliable as it combines the shape models such as active contours and active appearance model. The experiments prove that the approach is very effective.

At present, the large computation blocks the approach to be used in a real-time system, we will accelerate the algorithm to satisfy the real-time application.

Acknowledgments. The research work described in this paper was supported by 973 Basic Research Program of China (Project no. 2006cb504804).

References

1. Haritaoglu, S., Harwood, D., Davis, L.S.: W4: Real-Time Surveillance of People and Their Activities. IEEE Trans. On PAMI 22(8) (2000)
2. Siebel, N.T., Maybank, S.: Fusion of Multiple Tracking Algorithm for Robust People Tracking. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2353, pp. 373–387. Springer, Heidelberg (2002)
3. Radke, R.J., Andra, S., Al-Kofahi, O., Roysam, B.: Image Change Detection Algorithms: a Systematic Survey. IEEE Trans. on Image Processing 14(3), 294–307 (2005)
4. Zhao, T., Nevatia, R.: Bayesian Human Segmentation in Crowded Situations. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, vol. 2, pp. 18–20 (2003)

5. Chen, J., Shan, S., Yan, S., Chen, X., Gao, W.: Modification of the AdaBoost-based Detector for Partially Occluded Faces Track. In: Proceeding of International Conference on Pattern Recognition, vol. 2, pp. 516–519 (2006)
6. Dong, L., Parameswaran, V., Ramesh, V., Zoghlami, I.: Fast Crowd Segmentation Using Shape Indexing. In: Proc. IEEE Intl. Conf. on Computer Vision, pp. 1–8 (2007)
7. Wu, B., Nevatia, R.: Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors. In: Proc. Intl. Conference on Computer Vision, vol. 1, pp. 90–97 (2005)
8. Mohan, A., Papageorgiou, C., Poggio, T.: Example-based Object Detection in Image by Components. IEEE Trans. on PAMI 23(4) (2001)
9. Leibe, B., Seemann, E., Schiele, B.: Pedestrian Detection in Crowded Scenes. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 878–885 (2005)
10. Lienhart, R., Maydt, J.: An Extended Set of Haar-like Features for Rapid Object Detection. IEEE Conf. on International Conference on Image Processing 1, 900–903 (2002)
11. Collins, R., Lipton, A., Kanade, T.: Introduction to the Special Section on Video Surveillance. IEEE Trans. Pattern Anal. Machine Intell. 22(8), 745–746 (2000)
12. Wu, Y.N., Si, Z., Fleming, C., Zhu, S.C.: Deformable Template as Active Basis. In: Proc. IEEE Intl. Conf. on Computer Vision (2007)
13. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active Appearance Models. IEEE Transactions on Pattern Analysis and Machine Intelligence 23, 681–685 (2001)
14. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active Contour Models. International Journal of Computer Vision 1, 321–331 (1988)
15. Viola, P., Jones, M.J.: Rapid Object Detection using a Boosted Cascade of Simple Features. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 511–518 (2001)

A Geometric Active Contours Model for Multiple Objects Segmentation

Ning He[1], Peng Zhang[1], and Ke Lu[2]

[1] School of Mathematical Sciences, Capital Normal University, Beijing 100037, China
[2] College of Computing & Communication Engineering, Graduate University of Chinese Academy of Sciences, Beijing 100049, China
Hening_cnu@sina.com

Abstract. The technique of geometric active contours has become quite popular for a variety of applications, particularly image segmentation and classification problems. In traditional active contour models, snake initialization is performed manually by users, and topological changes, such as splitting of the snake, can not be automatically handled. In this paper, we present an automatic geometric active contours model which can extract multiple objects in an image without any manual assistance and completely eliminates the need of costly re-initialization procedure. The proposed framework is inspired by the geodesic active contour model and leads to a paradigm that is relatively free from the initial curve position. According to the proposed flow, the traditional boundary attraction term is replaced with a new force that guides the propagation to the object boundaries from both sides. This new geometric active contour model is implemented using a level set approach, thereby allowing dealing naturally with topological changes.

Keywords: Level sets, Geodesic active contour, Multiple objects segmentation.

1 Introduction

Active contours, or snakes [1] are a powerful technology to segment objects within images. There are two general types of active contours models: parametric active contours and geometric active contours. In this paper, we only concern the geometric active contours. Active contours can move under the influence of the internal forces, determined by the contour itself, and external forces derived from the image data or defined by user. The contour stops when its internal forces and its external forces bring into equilibrium. The internal forces compose of tensile forces and flexural forces. Each term has a weighted factor; we can satisfy all conditions in general by adjusting the weighted factor.

The external forces have various forms, such as the negative gradient of the edge derived from the image, and distance potential forces. These diverse forms can cause very different results. The traditional active contours models have several limitations that weaken their practicability of resolving the image segmentation problems. One of the limitations is the 'initialization sensitivity', that is, the initial contour must be close to the true boundary or it will likely converge to the wrong result. GVF (Gradient

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1141–1148, 2008.
© Springer-Verlag Berlin Heidelberg 2008

Vector Flow) [2] or its enhanced version GGVF (Generalized GVF) [3] is an effective method to solve this problem. GVF is a kind of external forces, which is computed as a diffusion of the gradient vectors of a gray-level or binary edge map derived from the image. GVF can expand the capture range remarkably, so the initial contour need not be as close to the true boundary as before. However, the proper initialization is still necessary, or else the active contour may converge to a wrong result.

In this paper, we propose a new boundary-based geometric flow for boundary extraction. This flow leads to the new external force that drives the propagated curves to the objects boundaries from either side. This new force is used to revise the geodesic active contour model, resulting on an elegant geometric flow. Furthermore, the implementation of this flow is done using level set methods where topological changes are naturally handled. All the processes are automatically completed; even drawing of a simple shape to initialize the contour is not needed, while it is necessary in other methods such as [4].

2 Background

2.1 Geodesic Active Contours

The geodesic active contour model [5] was introduced as a geometric alternative for snakes. It can be viewed as a classic snake since it overcomes some of their limitations. A traditional geodesic model is represented as a parametric curve $C = (x(s), y(s)), s \in [0,1]$, that moves through the spatial domain of an image to minimize the energy functional

$$E[C(p)] = \int_0^L g(|\nabla I(C(s))|)ds = \int_0^1 g(|\nabla I(C(s))|)\left|\frac{\partial C}{\partial p}(p)\right|dp \tag{1}$$

where g is a monotonically decreasing function

$$g : [0,+\infty] \to R^+, g(0) = 1, g(x) \to 0 \text{ as } x \to \infty$$

ds is the Euclidean arc-length element, and L is the Euclidean length of $C(p)$. In other words, when we try to detect an object we try to find the minimal length geodesic curve that best takes into account the desired image characteristics.

The minimization of the objective function is done using a gradient descent method. Therefore, the initial curve C_0 is deformed towards local minimum of $E(C(p))$ according to the following equation

$$C_t = g(|\nabla I|)KN - (\nabla g(\nabla I) \cdot N)N \tag{2}$$

where t denotes the time as the contour evolves, N is the inward Euclidean normal, and K is the Euclidean curvature. Each point of the contour should move along the normal direction in order to decrease the weighted length of C.

2.2 Gradient Vector Flow Active Contours

The Gradient Vector Flow (GVF) framework refers to the definition of a new elegant, bi-directional external force field [6] that captures the object boundaries from either

sides and can deal with concave regions. Opposite to this transform where in most of the cases a binary edge map is required, the GVF is estimated directly from the continuous gradient space.

The first required step within this framework is to determine a continuous edge-based information space which is given by a Gaussian edge detector

$$g(p)=\frac{1}{2\pi\sqrt{\sigma}}e^{-\frac{|\nabla(G_\sigma * I(p))|^2}{2\sigma^2}},\quad f(x,y)=1-g(p)$$

where $G_\sigma * I$ denotes the convolution of the input image with a Gaussian Kernel.

The gradient vector flow consists of a two dimensional vector field $\left[V(p)=(u(p),v(p)),p=(x,y)\right]$ that minimizes the following energy

$$E=\iint \mu\left(u_x^2+u_y^2+v_x^2+v_y^2\right)+|\nabla f|^2|V-\nabla f|^2\,dxdy \tag{3}$$

where μ is a blending parameter. Using the calculus of variations, it can be shown that the GVF can be found by solving the following Euler equations

$$\mathrm{v}_t=\mu\nabla^2 V-(V-\nabla f)|\nabla f|^2 \tag{4}$$

where v_t denotes the partial derivative of $\mathrm{v}(x,y,t)$ with respect to t, and ∇^2 is the Laplacian operator. Using GVF as the external forces greatly increases the capture range of snake, mainly influenced by the factor μ. Please refer to [2] and [3] for a thorough discussion of the GVF method. As the capture range expands, the GVF can pulls the contour to the desired boundary, even the initial contour is not close to the true boundary.

The rescaling of the Gradient Vector Flow field V leads to a new external force for the boundary term. This field contains mainly contextual information and the flow vectors of this field point to the "closest" object. However, it is a pity that the active contours models with GVF are still sensitive to the initialization. Fig. 1(a) shows that initialization of an active contour with a little circle which is eccentric to the desired boundary, i.e. the large circle. In Fig. 1(d) the contour has converged to an arc, it is apparently a wrong result. That is because the initial circle has not enclosed the 'center of divergence' of the GVF field.

To overcome above weakness we slightly modify the objective function as

$$E(V)=\iint \mu(u_x{}^2+u_y{}^2+v_x{}^2+v_y{}^2)+(1-\mu)|\nabla f|^2|V-\nabla f|^2\,dxdy \tag{5}$$

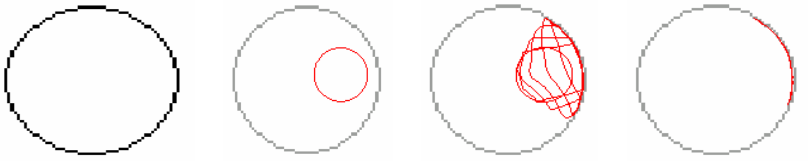


Fig. 1. (a) A circle (b) initial contour: a little circle which is eccentric to the large circle (c) contour deforms to an arc (d) final result

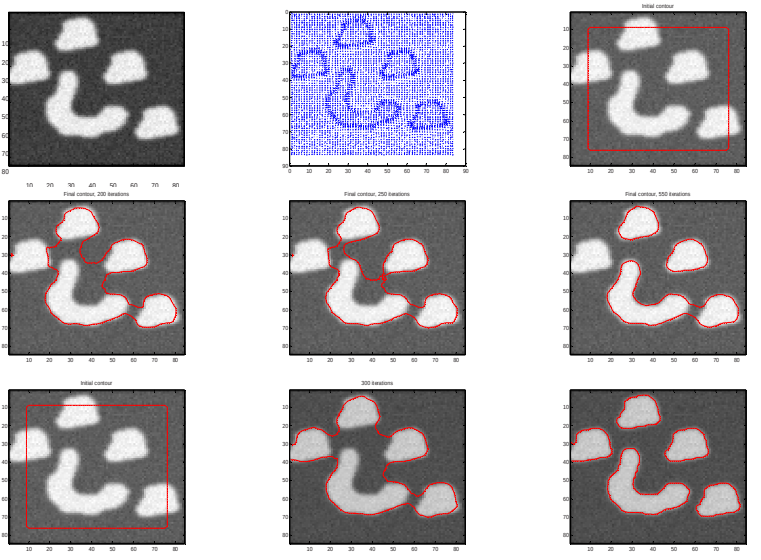


Fig. 2. Boundary extraction by the propagation of curves: Upper row: initial image, gradient vector flow and initial curve; Middle row: geodesic active contour and final result; Lower row: gradient vector flow geodesic active contour and final result

where μ is a diffusion constant parameter that balances the contribution between regularity and boundary attraction. The corresponding Euler equation is:

$$C_t = \mu gKN + (1-\mu)g(V \cdot N)N \tag{6}$$

It is quite clear that this flow under a constrained scenario, can produce an evolution similar to one obtained by geodesic active contour model. At the same time, when the propagating curve is located close to the object boundaries, then the flow has the same behavior with the second term of the geodesic active contour flow $[V \cdot N = \nabla g \cdot N]$. This limitation can be dealt with by incorporating an adaptive balloon force [7]. The modified model can be used to determine the direction of this new force, which is combined with the boundary force to define the boundary attraction term $[\mu gKN + (1-\mu)g]$. In this new flow the first force imposes a regularity constraint on the propagation and aims at shrinking the curve towards the object boundaries. The second force refers to a sophisticate boundary attraction term that moves the curve towards the objects boundaries from both sides (see Fig.2 Lower row).

3 New Geometric Active Contours Model

3.1 Level Set Formulation

Let $I : [0,a] \times [0,b] \to R^+$ be a given input in which the task of extracting the object contours is considered. The central idea of the level set [8] is to represent the moving curve $\partial C(t)$ as the zero-level set $\{\phi = 0\}$ of a function ϕ. This representation of $\partial C(t)$ is implicit, parameter-free and intrinsic. Additionally, it is topology free, since different

topologies of the zero level set correspond to the same topology of ϕ .It is easy to show, that if the embedding function ϕ deforms according to $\phi_t = F(p)|\nabla\phi(p,t)|$ .Then the corresponding moving front evolves according to: $C_t(p,t) = F(p)N(p)$, under the condition $[\phi(C(p,0),0) = 0]$.Thus, it can be shown that a steady state solution of the geodesic problem is given by:

$$\phi_t = g(|\nabla I|)K|\nabla\phi| + \nabla g(\nabla I)\cdot\nabla\phi \tag{7}$$

where $\phi(x,y,0) = \phi_0(x,y)$ and the curve C is represented by the zero level values of ϕ. The normal N, as well as the curvature K, can be estimated directly from the level set function ϕ $\left[N = -\frac{\nabla\phi}{|\nabla\phi|}, K = div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right]$. Now, the evolution of the proposed flows is equivalent to searching for a steady-state solution of the following equations:

$$\phi_t = \mu gK|\nabla\phi| - (1-\mu)g(V\cdot\nabla\phi) \tag{8}$$

3.2 Initialization of Level Set Function

Here we use the following functions as the initial function ϕ_0. Let Ω_0 be a subset in the image domain Ω, and $\partial\Omega_0$ be all the points on the boundaries of Ω_0, which can be efficiently identified by some simple morphological operations [9]. Then, the initial function ϕ_0 is defined as

$$\phi_0(x,y) = \begin{cases} -c, & (x,y) \in \Omega_0 - \partial\Omega_0 \\ 0, & (x,y) \in \Omega_0 \\ c & \Omega - \Omega_0 \end{cases} \tag{9}$$

where $c > 0$ is a constant.

3.3 Level Set Formulation with Penalizing Energy

In traditional level set methods, a common numerical scheme is to initialize the function ϕ as a signed distance function before the evolution, and then re-initialize the function ϕ to be a signed distance function periodically during the evolution. It is well known that a signed distance function must satisfy a desirable property of $|\nabla\phi| = 1$. Now we use the following integral $P(\phi) = \iint_\Omega (|\nabla\phi| - 1)^2 dxdy$ [9] as a metric to characterize how close a function ϕ is to a signed distance function in $\Omega \subset R^2$. The steepest descent process for minimization of the functional $P(\phi)$ is the gradient flow $\left[\Delta\phi - div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right]$. With the above defined functional $P(\phi)$ and proposed gradient vector flow geodesic active contour model, we propose a new formulation of geometric active contours model.

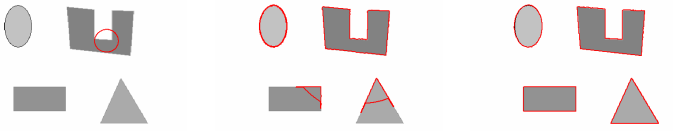
$$\phi_t = \lambda\left[\Delta\phi - div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right] + \mu div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)g|\nabla\phi| - (1-\mu)g(V\cdot\nabla\phi) \tag{10}$$

where $\lambda > 0$ is a parameter controlling the effect of penalizing the deviation of ϕ from a signed distance function, Δ is the Laplacian operator.

4 Experimental Results

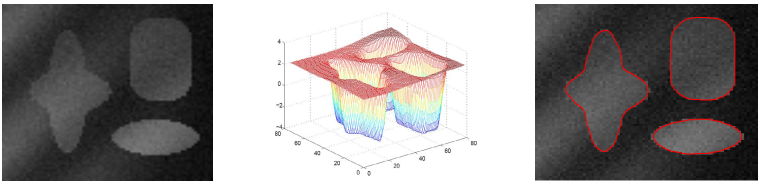
In this section, we apply the proposed segmentation method to three different images, and compare the results of the proposed method to the results of the GVF method.

Fig.3 shows the results using our method on an image with multiple objects which have different intensity and topology, at the same time we show that our method has less sensitivity to the initialization of the evolving contour. We used the arbitrarily initial contour as the initial level set function ϕ_0 is shown in Fig. 3(a), and evolve it using our evolution equation Eq.(10).



(a) (b) (c)

Fig. 3. Application to different intensity and topology image (a) Initial image and initial contour (b) The mediate result (c) The final result using our method



(a) (b) (c)

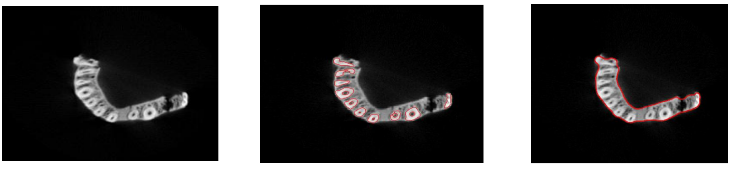
(d) (e) (f)

Fig. 4. (a) An input synthetic image with noisy (b) The normal velocity field in our method (c) Result obtained using our method (d) The GVF force field (e) Initial contour in the GVF method (f) Result obtained using the GVF method

Fig.4 shows the evolution of the contour on 83×65 pixel noisy synthetic image of three cells. As we can see, some parts of the boundaries of the cells are quite blurry. We use this image to demonstrate the robustness of our method in the presence of weak object boundaries and to show the results using our method and the GVF method. As shown in Fig. 4(c), the contour successfully evolved to the object boundaries, and the shapes of them are recovered very well.

We have compared our result with that obtained using the GVF method [10] on this image. Note that using the original GVF method on an image with multiple objects like this example, some parts of the moving contour may be parallel to the force field, thus cannot move further to find the objects. A balloon force was added in a modified GVF method in [2] when this deadlock happens. The force field of the GVF method for this image is shown in Fig. 4(d) and an initial contour as shown in Fig 4(e), the final result in Fig 4(f). We can see that the GVF method cannot provide a satisfactory result.

Fig.5 shows a segmentation of a CT image with multiple objects. Our method finds the objects accurately while the GVF method only find the outer boundary and cannot find the true boundaries of objects.



(a) (b) (c)

Fig. 5. Application to CT image (a) Initial image (b) Our Result (c) Result of GVF method

5 Conclusions

To summarize, in this paper, we have presented a geometric active contours model, which can automatically find multiple objects in an image. The boundary information is determined using a modified version of the gradient vector flow and is incorporated to the geodesic active contour model. This modification leads to a bi-directional flow that is relatively free from the initial curve conditions. While we introduce the energy penalizing function to eliminate the need of the re-initialization.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grant Nos. 60472071, 60532080, 60602062) and the National Science Foundation of Beijing (Grant No. 4051002).

References

1. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active Contour Models. Int. J. Comput. Vision 9, 321–331 (1988)
2. Xu, C., Prince, J.L.: Gradient Vector Flow: A new External Force for Snakes. In: IEEE Proc. Conf. on Comput. Vis. Patt. Recog. (CVPR), pp. 66–71 (1997)

3. Xu, C., Prince, J.L.: Generalized Gradient Vector Flow External Forces for Active Contours. Signal Processing 71, 131–139 (1998)
4. Choi, W., Lam, K., Siu, W.: An Adaptive Active Contour Model for Highly Regular Boundaries. Pattern Recognition 34, 323–331 (2001)
5. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. Int. J. Comput. Vision 22, 61–79 (1997)
6. Paragios, N., Mellina, O., Gottardo, O., Ramesh, V.: Gradient Vector Flow Fast Geodesic Active Contours. In: Proc. IEEE Int'l Conf. Comput. Vision, vol. I, pp. 67–73 (2001)
7. Paragios, N., Mellina, O., Gottardo, O., Ramesh, V.: Gradient Vector Flow Fast Geometric Active Contours. IEEE Trans. Pattern Anal.Machine Intell. 26, 402–407 (2004)
8. Osher, S., Fedkiw, R.: Level Set Methods. Technical Report CAM-00-08, Mathematics Department, UCLA (February 2000)
9. Li, C., Xu, C., Gui, C., Fox, M.D.: Level Set Evolution without Re-initialization: A New Variational Formulation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 430–436 (2005)
10. Xu, C., Prince, J.L.: Snakes, Shapes, and Gradient Vector Flow. IEEE Trans. Image Processing 7, 359–369 (1998)

A Mutation-Particle Swarm Algorithm for Error-Bounded Polygonal Approximation of Digital Curves

Bin Wang[1,⋆], Hua-Zhong Shu[1], Bao-Sheng Li[2] , and Zhi-Mei Niu[3]

[1] School of Computer Science and Engineering, Southeast University, Nanjing, 210096, P.R. China
binwang@seu.edu.cn
[2] Department of Radiation Oncology, Shandong Cancer Hospital, Jinan, 250117, P.R. China
[3] College of Computer Science and Engineering, Wuhan Institue of Technology, Wuhan, 430073, P.R. China

Abstract. This paper presents a particle swarm optimization algorithm (PSO) to solve error-bounded polygonal approximation of digital curves. Different from the existing PSO-based methods for polygonal approximation problem, the mutation operators borrowed from genetic algorithms, are incorporated into the PSO, so we call it MPSO. This scheme can increase the diversity of the population and help the particles effectively escape from the local optimum. Experiments were performed on three commonly used benchmark curves to test the effectiveness of the proposed MPSO. The results manifest that the proposed MPSO has the higher performance than the existing GA-based methods and PSO methods.

Keywords: Particle swarm optimization; Mutation operators; Polygonal approximation.

1 Introduction

Polygonal approximation is a good shape representation method which have the advantages of compactness and adjustable accuracy. There are two types of polygonal approximation problems that attract many researcher's interests[1]. One is to find an optimal polygon approximation with a minimum number of vertices on the pre-specified error tolerance. The other is to find an optimal polygon approximation with a minimal error on pre-specifying the number of polygon's vertices. In this paper, we focus on the former which is usually termed error-bounded polygonal approximation [2].

Polygonal approximation can be viewed as an combinatorial optimization problem. Some dynamic programming algorithms [3,4,5] were proposed to seek the global optimal polygonal approximation in the solution space. These methods

⋆ Corresponding author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1149–1155, 2008.
© Springer-Verlag Berlin Heidelberg 2008

often require a worst-case complexity of $O(N^4)$, where N is the number of curve points. Therefor, these methods are not suitable for real applications because the digital curves usually have a large amount of points. In recent years, there existed continuous interests of using nature-inspired algorithms, such as genetic algorithms (GA) [6,7,8,2,9,10], ant colony optimization (ACO)[11] and particle swarm optimization (PSO)[12], to solve the polygonal approximation problems. These methods have yielded encouraging results which bring us a promising way to solve polygonal approximation problems.

In this paper, we focus on using PSO algorithm to solve error-bounded polygonal approximation. The main contribution of our work is as follows: the existing PSO-based algorithm [12] may get trapped in local optimum because of the decrease of the diversity of the population on the search process, for overcoming this problem, we incorporate the mutation operator to the PSO algorithm, we term it MPSO, which can increase the diversity of the population and help the particles effectively escape from the local optimum. The experimental results show its higher performance than the existing GA-based methods and PSO methods.

This paper is organized as follows: Section 2 gives the definitions of error-bounded polygonal approximation. Section 3 reviews the existing PSO-based method. Section 4 presents the detail of the proposed MPSO. In section we present the experimental results and performance comparisons. Section 5 gives the conclusion.

2 Problem Formulation

A closed digital curve C can be represented by a sequence of points $C = \{p_1, p_2, \ldots, p_N\}$ which starts at point p_1 and continue to cross all the other points of curve in the clockwise direction along the curve and come back to the point p_1 finally.

Definition 1. *Let $\widehat{p_ip_j} = \{p_i, p_{i+1}, \ldots, p_j\}$ and $\overline{p_ip_j}$ represent the arc and the chord, respectively. The approximation error between $\widehat{p_ip_j}$ and $\overline{p_ip_j}$ is defined as*

$$e(\widehat{p_ip_j}, \overline{p_ip_j}) = \sum_{p_k \in \widehat{p_ip_j}} d^2(p_k, \overline{p_ip_j}), \tag{1}$$

where $d(p_k, \overline{p_ip_j})$ denotes the perpendicular distance from point p_k to the line segment $\overline{p_ip_j}$.

Definition 2. *The polygon V approximating the contour $C = \{p_1, p_2, \ldots, p_N\}$ is a set of ordered line segments $V = \{\overline{p_{t_1}p_{t_2}}, \overline{p_{t_2}p_{t_3}}, \ldots, \overline{p_{t_{M-1}}p_{t_M}}, \overline{p_{t_M}p_{t_1}}\}$, such that $t_1 < t_2 < \ldots < t_M$ and $\{p_{t_1}, p_{t_2}, \ldots, p_{t_M}\} \subseteq \{p_1, p_2, \ldots, p_N\}$, where M is the number of vertices of the polygon V. The approximation error between the curve C and its approximating polygon V is measured by integral square error (ISE) which is defined as*

$$ISE(V, C) = \sum_{i=1}^{M} e(\widehat{p_{t_i}p_{t_{i+1}}}, \overline{p_{t_i}p_{t_{i+1}}}), \tag{2}$$

Then the error-bounded polygonal approximation are formulated as follows:

Given the digital curve C and the error tolerance ε. Let Ω denotes the set of all the polygons which approximate the curve C. Let $\eta = \{V \mid V \in \Omega \wedge ISE(V, C) \leq \varepsilon\}$, Find a polygon $\alpha \in \eta$ such that

$$|\alpha| = \min_{V \in \eta} |V|, \tag{3}$$

where $|\alpha|$ denotes the cardinality of α.

3 Relative Work

The particle swarm optimization (PSO) is a kind of swarm intelligence algorithm and was developed by Eberhart and Kennedy [13]. PSO can be stated as follows. For a given optimization function $f(P)$, where P is a vector of n real-valued random variables, an initial swarm of particles is randomly generated. Each particle is represented as $P_i = (p_{i1}, p_{i2}, \ldots, p_{in})$, $i = 1, 2, \ldots, K$, where K is the swarm size. So, each particle represents a potential solution in the n-dimensional real number space and vector P_i represents the ith particle's position in the solution space. Except for the position, each particle P_i has another character: velocity vector $V_i = (v_{i1}, v_{i2}, \ldots, v_{in})$. The velocity V_i and position P_i are updated in each generation as follows:

$$v_{ij} = v_{ij} + c_1 r_{1j}(pbest_{ij} - p_{ij}) + c_2 r_{2j}(gbest_j - p_{ij}) \tag{4}$$

and

$$p_{ij} = p_{ij} + v_{ij} \tag{5}$$

where $pbest_i$ is the best previous position yielding the best fitness value for the ith particle; and $gbest$ is the best position discovered by the whole population. c_1 and c_2 are the acceleration constants and r_{1j} and r_{2j} are two random numbers in the range [0,1]. The velocity v_{ij} using Eq. 4 is restricted by a maximum threshold v_{max}.

PSO is initially developed for optimization problems in the continuous space. Eberhart and Kennedy [14] extended it to solve discrete combinatorial optimization problem. In the discrete version PSO, the Eq. 4 remains unchanged, except that it is constrained to the interval $[0.0, 1.0]$ by a sigmoid limiting transformation $S(v_{ij})$. The resulting change in position is then defined by the following rule: if $rand() < s(v_{ij})$, then $p_{ij} = 1$, otherwise $p_{ij} = 0$, where rand() is a random number in the range [0,1].

Yin [12] is the unique one who applied the above discrete version of PSO to solve error-bounded polygonal approximation problem. The method can be stated as follows. Each particle represented by a binary string is denoted as a approximating polygon. The length of the string is equal to the number of digital curve's points. Each position of the string corresponds to a point of the curve. If and only if its value is 1, its corresponding curve point is considered

as a vertex of the approximating polygon. For instance, given a curve $C = \{p_1, p_2, \ldots, p_{10}\}$ and a particle '1010100010'. Then the approximating polygon is $\{\overline{p_1p_3}, \overline{p_3p_5}, \overline{p_5p_9}, \overline{p_9p_1}\}$. The fitness of the particle is evaluated as follows: if $E > \varepsilon$ then $fitness = -E/\varepsilon N$, otherwise $fitness = 1/\sum_{j=1}^{N} p_{ij}$. Yin sets the parameter of algorithms as follows: the number of particles is 20, acceleration constants $c_1 = c_2 = 2$ and the maximal velocity $v_{max} = 6$.

4 The Proposed MPSO

To increase the diversity of the population and help the particle to escape from the local optimum, we apply the mutation scheme to the above Yin's PSO-based method. In MPSO, each particle will suffer from a mutation operator. Two kind of mutation operators, single-point mutation and gene-shift mutation, are developed for MPSO. The single-point mutation is illustrated as follows: among the genes whose value are 1 on the particle, randomly select one and change its gene value to 0, this operator corresponds to remove a vertex from the polygon. Another mutation operator: gene-shift mutation is described as follows: among the genes whose value are '1' on the particle, randomly select one and shift it a site to left or right and set 0 to the original gene site. The algorithm flow of MPSO is as follows.

input the digital curve $C = \{p_1, p_2, \ldots, p_N\}$, the error tolerance ε.
output the polygon U which approximates C.
step 1 Generate K particles $P_1, P_2, \ldots, P_k$.
step 2 Calculate the fitness value of each particle.
step 3 Determine $Pbest_i$ for each particle P_i and $gbest_i$ for the whole swarm.
step 4 Update velocities v_{ij} according to Eq. 4 retricted by a maximum threhold .
step 5 Set $S(v_{ij}) = 1/(1 + e^{-v_{ij}})$.
step 6 if $rand() \leq Sv_{ij}$, then $P_{ij} = 1$, otherwise $P_{ij} = 0$.
step 7 For each particle, perform a mutation operator against it. The way of selecting which kind of mutation operator is as follows: if $rand() \leq 0.5$, then perform single-point mutation, otherwise perform gene-shift muation.
step 8 update the generation number g=g+1. If g achieve the pre-specified number of iterations, then go to step 2, otherwise output U.

5 Experimental Results and Discussion

To evaluate the performance of the proposed MPSO, Three benchmark curves, leaf curve, chromosome curve and semicircle curve (see Fig. 2 (a),(f),(k), which have been widely used in literatures [6,7,11,15] are utilized in the experiments. Their chain codes can be obtained in [15]. Among these curves, (a) is a leaf-shaped curve, (f) is a chromosome-shaped curve, and (k) is a curve with four semi-circles The number of their curve points is 120, 60 and 102, respectively.

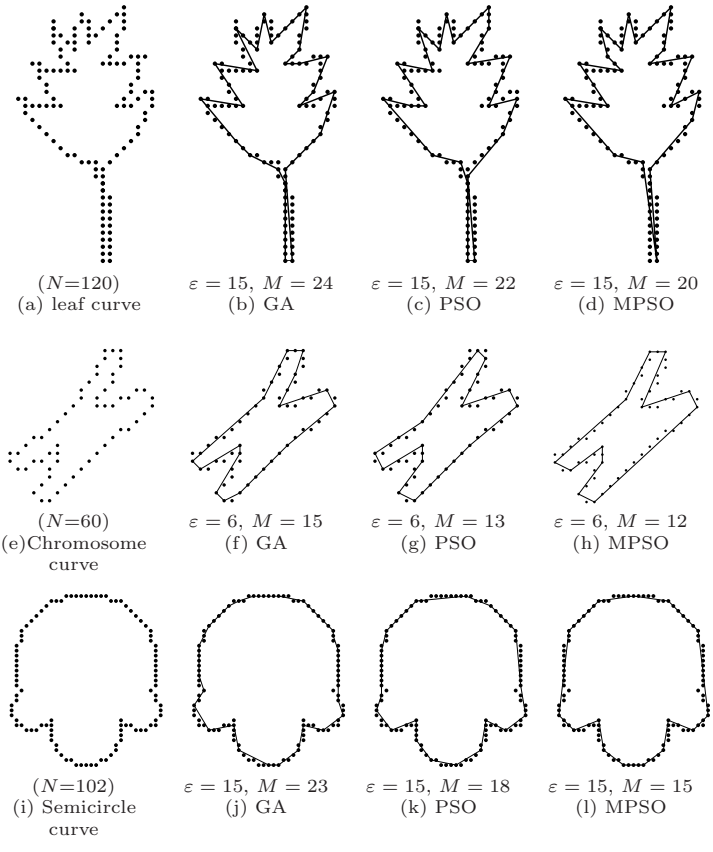


Fig. 1. Three benchmark curves and the comparative results of GA, PSO and MPSO, where N, ε and M denote the number of points on the curve, the error tolerance and the number of vertices of the polygon approximating the curve, respectively

The other two algorithms, genetic algorithms (GA)[6] and particle swarm optimization (PSO) [12], are used as the comparison objects. The parameters for MPSO is set as follows: the number of particles is set to 20; the acceleration constants $c_1 = c_2 = 2$; the maximal velocity $v_{max} = 6$ and the number of iterations is set to 500. The parameters for the other competing methods are set as the ones provided by these literatures. Since these algorithms are all probabilistic search methods, each competing methods will be conducted ten times, the best results and the averaging results of these runs will be reported. The experimental platform is a PC with a pentium-4 2.4GHz CPU running Windows XP and all the competing methods are coded in Borland Delphi 7.0.

The simulation results, including the average results M, the standard deviation σ and the average time t over ten independent runs are reported in table 1. Fig. 2 shows the finally obtained approximation polygons with the number of

Table 1. Eperimental results of three benchmarks curves for GA, PSO and MPSO

Curves	ε	GA		PSO		MPSO	
		$M(\sigma)$	t	$M(\sigma)$	t	$M(\sigma)$	t
	150	15.6(0.6)	0.13	11.8(1.5)	0.24	10.1(0.5)	0.25
	100	16.3(0.5)	0.13	14.3(2.0)	0.25	12.5(0.3)	0.26
Leaf	90	17.3(0.5)	0.13	14.9(2.5)	0.25	12.8(0.2)	0.25
($N = 120$)	30	20.5(0.6)	0.13	20.2(1.9)	0.25	17.2(0.4)	0.26
	15	23.8(0.6)	0.13	25.7(2.4)	0.25	21.5(0.7)	0.26
	30	7.3(0.4)	0.07	7.3(0.9)	0.13	6.0(0.0)	0.13
	20	9.0(0.6)	0.07	9.3(1.3)	0.13	7.4(0.5)	0.13
Chromosome	10	10.2(0.4)	0.07	12.0(1.5)	0.13	10.0(0.0)	0.13
($N = 60$)	8	12.2(0.5)	0.07	13.0(1.1)	0.13	11.5(0.3)	0.13
	6	15.2(0.6)	0.07	14.9(1.4)	0.13	12.9(0.7)	0.13
	60	13.2(0.4)	0.11	12.6(0.5)	0.21	10.5(0.5)	0.21
	30	13.9(0.7)	0.11	15.1(1.2)	0.21	13.6(0.3)	0.21
Semicirle	25	16.8(0.7)	0.11	16.6(1.6)	0.21	14.3(0.5)	0.21
($N = 102$)	20	19.2(0.6)	0.11	17.8(1.8)	0.22	15.9(0.3)	0.22
	15	23.0(0.9)	0.11	21.3(2.2)	0.22	15.9(0.7)	0.22

vertices (M) under the pre-specified error tolerance (ε) for visual comparison of the approximation, where σ and t are the standard deviation of solutions and the average computation time (in seconds), respectively. The simulation results show that for the same testing curve, under the same pre-specified error tolerance, MPSO produces approximating polygon with fewer number of vertices than the other approaches and its standard deviation of the solutions obtained is also smaller than the other ones which shows its higher robustness. It is noted that from the table 1, the computational cost for MPSO and PSO are comparative, which shows that the computational cost for the additional mutation operator can be neglected. Although genetic algorithm has a lower computational cost, the quality of the solution is worse than the proposed MPSO.

6 Conclusion

A mutation-PSO algorithm has been proposed to solve the error-bounded polygonal approximation. Different from the existing PSO-based method, the mutation operators borrowed from genetic algorithms, are incorporated into the PSO, so we call it MPSO. This scheme can increase the diversity of the population and help the particles effectively escape from the local optimum. Experiments has been performed on three commonly used benchmark curves to test the effectiveness of the proposed MPSO. The results manifest that the proposed MPSO has the higher performance than the existing GA-based method and PSO method.

Acknowledgements. The research work presented in this paper is supported by National Natural Science Foundation of China, project No.30670617.

References

1. Imai, H., Iri, M.: Polygonal Approximation of a Curve (Formulations and Algorithms. In: Toussaint, G.T. (ed.) Computational Morphology, pp. 71–86. North-Holland, Amsterdam (1988)
2. Sun, Y.N., Huang, S.C.: Genetic Algorithms for Error-bounded Polygonal Approximation. Int.J. Pattern Recognition and Artificial Intelligence. 14, 297–314 (2000)
3. Dunham, J.G.: Optimum Uniform Piecewise Linear Approximation of Planar Curves. IEEE Transactions on pattern Analysis and Machine Intelligence 8, 67–75 (1986)
4. Sato, Y.: Piecewise Linear Approxiamtion of Planes by Perimeter Optimization. Pattern Recognition 25, 1535–1543 (1992)
5. Perez, J.C., Vidal, E.: Optimum Polygonal Approximation of Digitized Curves. Pattern Recognition Letter 15, 743–750 (1994)
6. Yin, P.Y.: Genetic Algorithms for Polygonal Approximation of Gigital Curves. Int. J. Pattern Recognition Artif. Intell. 13, 1–22 (1999)
7. Yin, P.Y.: A New Method for Polygonal Approximation Using Genetic Algorithms. Pattern Recognition Letter. 19, 1017–1026 (1998)
8. Huang, S.C., Sun, Y.N.: Polygonal Approximation Using Genetic Algorithms. Pattern Recognition 32, 1409–1420 (1999)
9. Ho, S.Y., Chen, Y.C.: An Efficient Evolutionary Algorithm for Accurate Polygonal Approximation. Pattern Recognition 34, 2305–2317 (2001)
10. Sarkar, B., Singh, L.K., Sarkar, D.: A Genetic Algorithm-based Approach for Detection of Significant Vertices for Polygonal Approximation of Digital Curves. International Journal of Image and Graphics 4, 223–239 (2004)
11. Yin, P.Y.: Ant Colony Search Algorithms for Optimal Polygonal Approximation of Plane Curves. Pattern Recognition 36, 1783–1997 (2003)
12. Yin, P.Y.: A Discrete Particle Swarm Algorithm for Optimal Polygonal Proximation of Digital Curves. Journal of Visual Communication and Image Representation 15, 241–260 (2004)
13. Eberhart, R.C., Kennedy, J.: A New Optimizer Using Particle Swarm Theory. In: Proc. 6th Symp. Micro Machine and Human Science, Nagoya, Japan, pp. 39–43 (1995)
14. Eberhart, R.C., Kennedy, J.: A Discrete Binary Version of the Particle Swarm Algorithm. In: IEEE International Cofference on SMC, pp. 4104–4108 (1997)
15. Teh, H.C., Chin, R.T.: On Detection of Dominant Points on Digital Curves. IEEE Trans. Pattern Anal. Mach. Intell. 11, 859–872 (1989)

Dimensionality Estimation for Self-Organizing Map by Using Spectral Clustering

Naoyuki Tsuruta[1], Saleh K.H. Aly[2], and Sakashi Maeda[1]

[1] Department of Electronics Engineering and Computer Science, Fukuoka University
8-19-1 Nanakuma Jonan-ku, Fukuoka, 814-0180, Japan
[2] Department of Intelligent Systems , Kyushu University, Japan
{tsuruta,aly,maeda}@mdmail.tl.fukuoka-u.ac.jp
http://www.tl.fukuoka-u.ac.jp/~tsuruta/

Abstract. Tasks of image recognition become important components for multimodal man-machine interface. For developing feasible components, problems of huge dimensionality and non-linearity must be resolved. We have been applying Self-Organizing Map (SOM) to feature representation stage of lip-reading and face recognition, and have appealed the advantages of SOM for dimensionality reduction and nonlinear feature representation. However, tasks of trial and error to decide the appropriate dimensionality of SOM can be a difficulty for developing image recognition. For this problem, we propose a dimensionality estimation method for SOM by using spectral clustering (SC). SC also has a characteristic of non-linear topographic mapping, and its fruitage suggests the dimensionality of feature space. In the section of experimental results, we will show relations between estimated dimensionalities of SOM and total recognition accuracies. The results emphasize feasibility of this proposed method.

Keywords: Dimensionality estimation, Self-organizing Map, Spectral Clustering, feature extraction, image recognition.

1 Introduction

Image recognition has become a key technology for multi-modal man-machine interface, and has applied to important components, such as lip-reading, face recognition[7], facial emotion recognition[6] and gesture recognition[8]. For developing feasible components, in feature extraction stage, problems of huge dimensionality and non-linearity[5] must be resolved. However, linear methods, such as principle component analysis (PCA)[4] and linear discriminant method (LDM), are commonly used yet. Self-organizing feature map (SOM)[3] and spectral clustering (SC)[9,10] are candidates of the non-linear feature extraction. SOM and SC are shortly discribed in Section 3 and 4, respectively.

We have been applying SOM to feature representation stage of lip-reading and face recognition, and have appealed the advantages of SOM for dimensionality

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1156–1163, 2008.
© Springer-Verlag Berlin Heidelberg 2008

reduction and nonlinear feature representation. However, tasks of trial and error to decide the appropriate dimensionality of SOM can be a difficulty for developing image recognition. For this problem, we propose a dimensionality estimation method for SOM by using SC. SC also has a characteristic of non-linear topographic mapping, and its fruitage suggests the dimensionality of feature space. In Section 5, we will show experimental results. Estimated dimensionalities of SOM by using SC give good total recognition accuracies. The results emphasize feasibility of this proposed method.

2 Overview of Image Recognition Process

2.1 Basic Architecture

Image recognition consists of three stages:

1. calibration or normalization stage,
2. feature extraction or representation stage,
3. recognition stage.

For recognition stage, state of the art methods including nonlinear methods were proposed. On the other hand, linear methods, such as principle component analysis and linear discriminant method, are commonly used yet for feature extraction stage. To overcome the problem of huge dimentionality and nonlinearity, we need a new method for the feature extraction stage.

2.2 Roles of Feature Extraction Stage

The most important role of feature extraction stage for image recognition is reduction of dimensionality, because huge dimensionality of image causes problems of computational complexity and Hughes phenomenon. To overcome these problem, linear reduction methods such as PCA and LDM are traditionaly used.

Another important role is parameterizaion of feature space. Fig. 1 shows a feature space and distribution of training data for a lip-reading system[2]. The training data consists of 2000 lip images, each of which has 160×120 pixels. The feature space is expanded by the first and the third principal component of the training data. It is clear that distribution of lip images is 'winding,' and that the distance shouldn't be measured linearly but should measure along the winding Geodesic on the manifold instead. In that case, there is a possibility that dimensionality of Geodesic is much lower than one of original space.

3 Self-Organizing Feature Map (SOM)

SOM is one of the most widely used artificial neural networks, which uses unsupervised competitive learning. It is a mathematical model that simulates the functioning of the hyper column in human brain. In this section, assuming two dimensional SOM, we show a traditional description for SOM[1]. Units, usually

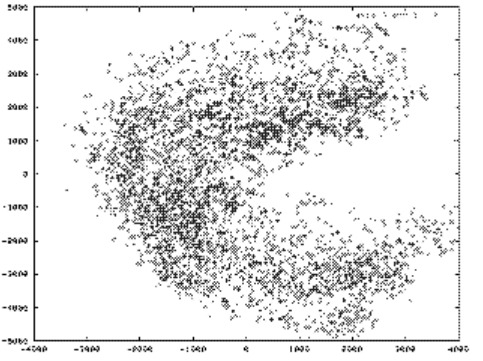


Fig. 1. A distribution of 2000 lip images in a feature space expanded by the 1st and 3rd principal components

called neurons, are tuned to different input vectors $\mathbf{I}$. A reference vector $\mathbf{w}_i$ is associated with each neuron i. In each training step, one sample input vector $\mathbf{I}$ is considered and a similarity measure, usually taken as the Euclidian distance, is calculated between the input sample and all reference vectors of neuron. The best matching neuron c is selected as the neuron that satisfies

$$||\mathbf{I} - \mathbf{w}_c|| = \min_i(||\mathbf{I} - \mathbf{w}_i||), \tag{1}$$

and called "winner."

After selecting the winner neuron, all the reference vectors are updated according to the rule

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{ci}(t)\left(\mathbf{I} - \mathbf{w}_i(t)\right), \tag{2}$$

where

$$h_{ci}(t) = \alpha(t)\exp\left(\frac{||r_c - r_i||}{2\sigma^2(t)}\right). \tag{3}$$

h_{ci} is the neighborhood kernel around the winner at time t. $\alpha(t)$ is the learning rate $(0 < \alpha(t) < 1)$ and is decreased gradually toward zero and $\sigma^2(t)$ is a factor used to control the neighborhood kernel. r_c and r_i are the location vectors of neuron c and i, respectively, in the two dimensional SOM array. The term $||r_c - r_i||$ is referring to the distance between the neurons. After the training data are exhausted, the neuron sheet is automatically organized into a meaningful two-dimensional order denoted by feature map (or reference vector).

The SOM reference vector has the two following remarkable characteristics:

- The Probability Distribution Function (PDF) of the reference vector is a good approximation for the PDF of the training data.
- The topographical order of the training data is preserved in the reference vector, even if the dimensionality of the SOM is smaller than that of the training data.

4 Spectral Clustering (SC)

SC is proposed from graph theory field. Now, consider a minimum cut set problem of non-directional graph. Let an array $\mathbf{W} = [w_{ij}]$ denote connectivity between node i and node j. If node i and j have a connection, then $w_{ij} = 1$, otherwise $w_{ij} = 0$. Additionally, new values, which indicate that each node belongs to which of graph A or graph B, are introduced, as following:

$$q_i = \begin{cases} 1 & \text{if } i \in A \\ -1 & \text{if } i \in B \end{cases} \tag{4}$$

Then, cut size of the graph is described by the following equation:

$$\begin{aligned} J &= CutSize \\ &= \frac{1}{4}\sum_{i,j} w_{ij}(q_i - q_j)^2 \\ &= \frac{1}{4}\sum_{i,j} w_{ij}(q_i^2 + q_j^2 - 2q_i q_j) \\ &= \frac{1}{2}\sum_{i,j} q_i(d_i\delta_{ij} - w_{ij})q_j \\ &= \frac{1}{2}\mathbf{q}^T(\mathbf{D} - \mathbf{W})\mathbf{q}, \end{aligned} \tag{5}$$

where D is a diagonal matrix and δ_{ij} is Kronecker function:

$$D = [\mathrm{diag}(d_i)], \tag{6}$$

$$d_i = \sum_j w_{ij}, \tag{7}$$

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{8}$$

By relaxing q_i from binary to real, $\min J(\mathbf{q})$ is given as an eigen vector of

$$(\mathbf{D} - \mathbf{W})\mathbf{q} = \lambda\mathbf{q}. \tag{9}$$

Because elements of $\mathbf{q}$ are already relaxed to real value, q_i takes $-1 \leq q_i \leq 1$. Therefore, nodes are reordered by q_i. There are many modified version of Eqn. (9). In this paper, we use

$$(\mathbf{D} - \mathbf{W})\mathbf{q} = \lambda\mathbf{D}\mathbf{q}, \tag{10}$$

because of our experiences.

The eigen value takes $0 \leq \lambda \leq 1$. The eigen vector for the minimum eigen value is optimal solution, because we are going to minimize the cut size J. The first eigen value, however, is zero, and its eigen vector is a direct current component. Meaningful solution is the second eigen vector or later than second. Element's

value q_i indicatres topographic order: an order to aligne the data along with non-linear principal curve. It is well known that SC's behavior on preserving topographic order is very similar with SOM's one.

In the proposed method, w_{ij} are calculated from a pair of training images, $\mathbf{I_i}$ and $\mathbf{I_j}$, by the following equations.

$$w_{ij} = \exp\left(-\frac{s_{ij}}{2\sigma}\right), \tag{11}$$

$$s_{ij} = ||\mathbf{I}_i - \mathbf{I}_j||^2, \tag{12}$$

$$\sigma = E(s_{ij}), \tag{13}$$

where $E(s_{ij})$ is mean of s_{ij} for all combination of i and j. In the solution of Eqn. (10), eigen values decide an order of principal curve with the same manner of PCA. And elements q_i corresponding to selected eigen values expand an image feature space. We put an interest on the eigen value in the next section. As unlike PCA, eigen value is not measurement of statistic parameter such a variance, but it can be an indicator to decide dimensionality of SOM.

5 Experimental Results

5.1 Lip-Reading

Our lip-reading system[2] inputted a sequence of images for a phrase consists of a few words, and outputted a phrase as a recognition result. The system consisted of a SOM/FDCT (Fast Discrete Cosine Transform) for feature representation, and a Hidden Malkov Model recognizer. The SOM selected the nearest neighbor neuron for each image. The index of the winner neuron was a feature value. A sequence of the feature values was an input for HMM recognizer. We developed a lip-reading systems for Arabic.

Table 1 and Fig. 2 show a relation between dimensionality of subspace generated by FDCT's coefficients and recognition accuracy[11]. CP means "cumulative proportion." For lower dimensional subspace, when number of co-efficiencies increases, accuracy and CP also increase. However, accuracy saturates from around 9 dimensional case, and decrease before CP becomes near to 100%. This phenomenon is well known as "Hughes phenomenon." This result shows that CP cannot predict Hughes phenomenon and cannot be a measurement for estimation of appropliate dimensionality.

Table 1. A relation of dimensionality, recognition accuracy and CP using subspace expanded by FDCT coefficients for images of Arabic speech

Dimensionality	2	4	5	7	8	9	11
Word	43.6	57.7	66.7	67.9	**69.2**	67.9	63.5
Phrase	11.2	25.9	40.7	40.7	**45.1**	40.7	36.4
CP(%)	46.2	58.7	60.5	64.0	**65.4**	66.9	69.2

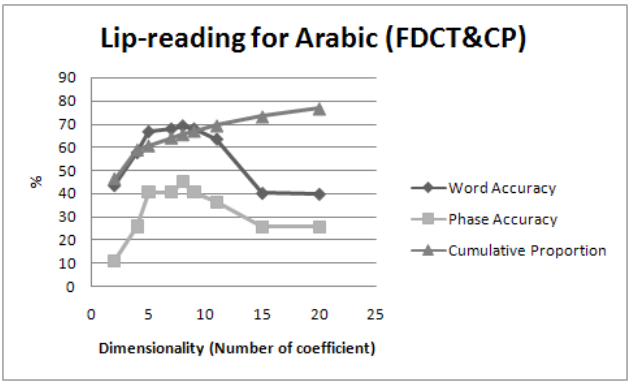


Fig. 2. A relation of dimensionality, recognition accuracy and CP for Arabic lip-reading by using FDCT coefficients

Table 2. A relation of recognition accuracy, eigen value of SC and dimentionality of SOM for Arabic lip-reading

Dimensionality	1	2	3	4	5
Word		66.7	70.5	76.9	77.7
Phrase		40.7	48.7	51.8	52.5
Eigen value	0.897	0.943	0.959	0.973	0.978

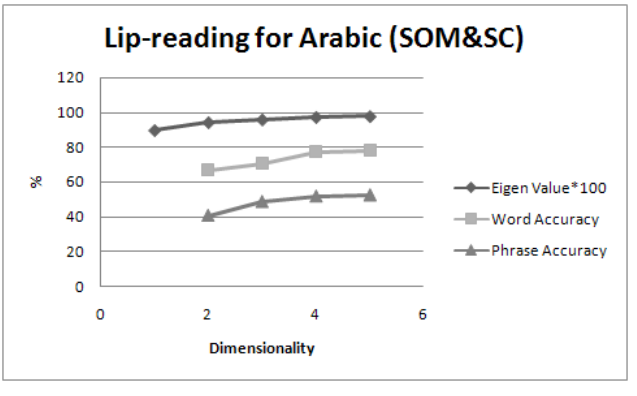


Fig. 3. A relation of recognition accuracy, eigen value of SC and dimentionality of SOM for Arabic lip-reading

On the other hand, Table 2 and Fig. 3 conclude the relation between dimensionality of the desired feature map, eigen value of SC and recognition accuracy[1]. Fig. 2 clearly shows that eigen value and recognition accuracy form very similar

[1] For dimensionalities higher than six, SOM required huge number of neurons and was hard to be trained.

Table 3. A relation of eigen value of SC, recognition accuracy and dimentionality of SOM for face recognition

Dimensionality	1	2	3	4	5	6	7
Accuracy		20.7	33.6	43.0	47.3	49.1	49.4
Eigen value	0.797	0.912	0.959	0.975	0.984	0.986	0.989

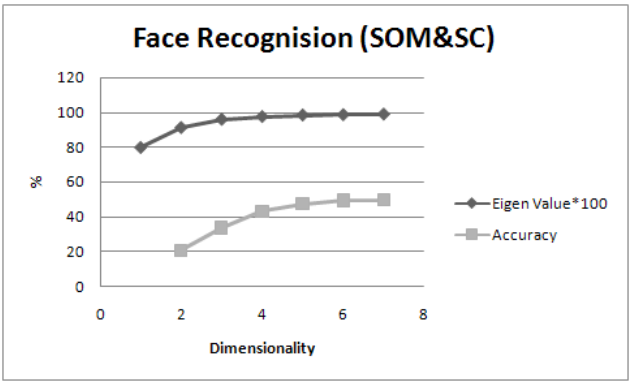


Fig. 4. A relation of eigen value of SC, recognition accuracy and dimentionality of SOM for face recognition

curves for each other, and saturate from the almostly same dimensionality. This result means that eigen value of SC can suggest the start point, from where recognition accuracy begins to saturate.

5.2 Face Recognition

We used Yale B face database[12]. Number of images in this database is 20,520 images; half of them are reserved for training and the other half are reserved for recognition. We used a SOM as a feature extractor and a support vector machine (SVM) as a feature recognizer.

Table 3 and Fig. 4 conclude the relation between dimensionality of the desired feature map, eigen value of SC and recognition accuracy [2] . Fig. 4 also shows that eigen value and recognition accuracy also form very similar curves for each other, and saturate from the almostly same dimensionality. This result also means that eigen value of SC can suggest the start point, from where recognition accuracy begins to saturate.

[2] For dimensionalities higher than eight, SOM required huge number of neurons and was hard to be trained.

6 Conclusion

In this paper, we proposed Spectral Clustering base dimensionality estimation method for SOM. Eigen values of Spectral Clustering solution is not measurement of statistic parameter such that variance, but we showed that it could be an indicator to decide dimensionality of SOM. The estimated dimensionality for each application: lip-reading for Arabic and face recognition, were quite match with the best case of our conventional experimental results.

References

1. Kohonen, T.: Self-Organizing Maps. Second Edition. Springer, Heidelberg (1997)
2. Sagheer, A., Tsuruta, N., Taniguchi, R., Maeda, S.: Visual Speech Features Representation for Automatic Lip-Reading. In: The 30th IEEE Int. Conf.on Acoustics, Speech and Signal Processing (ICASSP 2005), vol. 2, pp. 781–784 (2005)
3. Kohonen, T.: The Self Organizing Map. Neurocomputing 21, 1–6 (1998)
4. Jolliffe, I.: Principal Component Analysis, 2nd edn. Springer, Heidelberg (2002)
5. Lu, B., Hsieh, W.: Simplified nonlinear principal component analysis. In: Int. Conf. On Neural Networks (IJCNN 2003), vol. 1, pp. 759–763 (2003)
6. Sim, T., Baker, S., Bsat, M.: The CMU pose, illumination and expression database. The IEEE Trans. on Pattern Analysis and Machine Intelligence 25(12), 1615–1618 (2003)
7. Stan, Z., Jain, A.K. (eds.): Handbook of Face Recognition, 1st edn. Springer, Heidelberg (2005)
8. Tobely, T.E., Tsuruta, N., Amamiya, M.: On-line Hand Gesture Recognition Using the Hypercolumn Neural Network. In: Int. Conf. on Artificial Intelligence and Applications, pp. 198–203 (2003)
9. Pothen, A., Simon, L.H.K.: Partitioning sparse matrices with eigenvectors of graphs. SIAM J. Matrix Anal. 11(3), 430–452 (1990)
10. Ding, C.H.Q.: `http://crd.lbl.gov/~cding/`
11. Sagheer, A., Tsuruta, N., Taniguchi, R., Maeda, S.: HyperColumn Model vs. Fast DCT for Feature Extraction in Visual Arabic Speech Recognition. In: IEEE Int. Symposium on Signal Processing and Information Technology, pp. 761–766 (2005)
12. Sagheer, A.: Self Organizing Neural Networks for High Dimensional Feature Space Problem in Pattern Recognition Applications, Ph.D thesis in Kyushu University, Japan (2007)

Micromanipulation Using a Microassembly Workstation with Vision and Force Sensing

Hakan Bilen and Mustafa Unel

Faculty of Engineering and Natural Sciences, Sabanci University
34956 Istanbul, Turkey
hakanbil@su.sabanciuniv.edu
munel@sabanciuniv.edu

Abstract. This paper reports our ongoing work on a microassembly workstation developed for efficient and robust 3D automated assembly of microobjects. The workstation consists of multiple view imaging system, two 3-DOF high precision micromanipulators, and a 3-DOF positioning stage with high resolution rotation control, force sensing probe and gripper, and the control software system. A hybrid control scheme using both vision and force sensory information is proposed for precise and dexterous manipulation of microobjects. A micromanipulation experiment that aims to locate the microspheres to the predefined configuration by using an integrated vision and force control scheme is successfully demonstrated to show the validity of the proposed methods.

Keywords: microassembly, micromanipulation, visual servoing, force control.

1 Introduction

In recent years with the advances in the fields of micro- and nano-technology, handling micro-scale objects and operating in micro and submicron domains have become critical issues. Although many of commercially available micro devices such as inkjet printer heads and optical components are currently produced in monolithic fabrication process with micromachining, other products such as read/write heads for hard disks, fiber optics assemblies and RF switches/filters require more flexible assembly process. In addition, many biological micromanipulations such as invitro-fertilization, cell characterization and treatment are currently performed manually. Requirement of high-precision, repeatable and financially viable operations in these tasks has given rise to the elimination of direct human involvement, and autonomy in micromanipulation and microassembly.

In the literature, several research efforts on autonomous micro-manipulation and assembly tasks can be found in the areas of microelectromechanical (MEMS) assembly systems [1]-[3] and microrobotic cell injection systems [4]-[6]. Kim et al. [1] proposed a hybrid assembly method which combines the vision-based microassembly and the scaled teleoperated microassembly with force feedback. With these tools, 980 nm pump laser is manufactured by manipulating and assembling opto-electrical components. Dionnet et al [7] designed a system which

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1164–1172, 2008.
© Springer-Verlag Berlin Heidelberg 2008

uses vision and force feedback techniques that allow to achieve accurate and safe autonomous micromanipulation. In the system, microparticles are positioned by pick-up using adhesion forces and release by rolling. Wang et al [4] presents a microrobotic system for fully automated zebrafish embryo injection based on computer vision and motion control. The microrobotic system performs autonomous injection at a high speed, with a successful survival rate. In spite of the numerous works in this area, only a few works [9] and [10] reported fully automatic microassembly tasks in 3D under optical microscopes.

In this work, an open architecture microassembly workstation in which fully automated 3D micromanipulation is executed under vision and force control is proposed. A hybrid control scheme using vision and force is developed to achieve both high precision and dexterity in micromamanipulation. The vision subsystem enables the system to move the end effector to the desired position based on the real-time detection of microobjects and probe tip. The force feedback is used to dexterously manipulate the microparticles without losing contact or applying excessive forces.

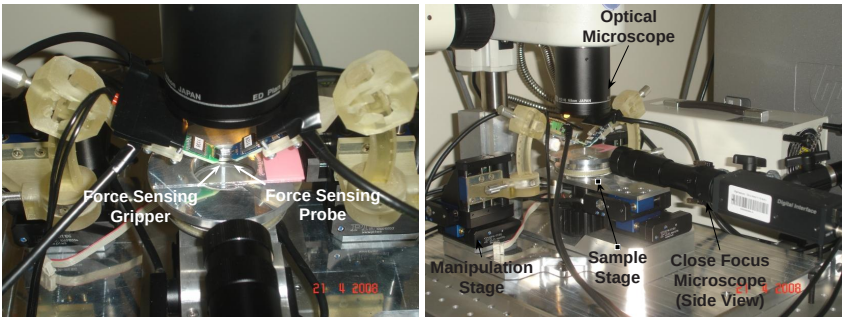
This paper is organized as follows: Section 2 presents the hardware and software components of the workstation. Section 3 explains the system calibration. Section 4 is on micromanipulation using vision and force. A micromanipulation experiment is also provided in this section. Finally, Section 5 concludes the paper.

2 Microassembly Workstation

A microassembly workstation is designed for 3D manipulation and assembly of microobjects. The workstation uses two types of end effectors, a force sensing probe and a microgripper which have 0.3 μN resolution at 30Hz, shown in Fig. 1.(a). The force sensing probe and gripper are mounted on two separate 3-DOF fine positioning stages (50 nm closed loop precision). On an x-y-θ positioning stage, (50 nm close loop precision and 4.5×10^{-5} degrees rotation resolution) a glass slide is mounted and is positioned under the force sensing probe and microgripper, depicted in Fig. 1.(b). On the glass slide, the samples, polystyrene balls and biological cells which are used in experiments can be located.

The samples are placed under a stereo optical microscope, Nikon SMZ1500 with 1.5x objective and 0.75 : 11.25 zoom ratio. The optical microscope is coupled with two CCD cameras. While a CCD camera provides coarse view for the sample stage, another one is used for fine and narrower field of view. In addition, a side camera coupled to a 35x close focus microscope with 63 mm working distance and 7.2x7.2 mm^2 field of view is employed to provide a side view to acquire the height information between the sample stage and the end effector. A workstation is employed to acquire the images from the cameras and process the visual information. The resolution of the coarse, fine and side cameras are approximately 4, 1.3 and 2.5 pixels/μm for 2x zoom level respectively.

The positioning stages are controlled by a dSpace ds1005 motion control board. The control software is written in the C programming environment and



(a) (b)

Fig. 1. Microassembly Workstation, (a) microgripper and probe, (b) sample and manipulator stages

provides real-time control for moving and positioning the stages precisely. The computer vision software is written in C++ language by using OpenCV libraries.

The optical microscope can provide only 2D visual information of the x-y plane due to small depth of field. Although exploiting defocus can yield a coarse information along the z axes, it results in poor accuracy for micron precision applications and it is computationally expensive. Therefore, a close focus microscope is used to monitor the image information in x-z plane and interaction between the sample stage and the end effector. Since it has a relatively large working distance, the samples and the probe/gripper can be well focused for an adequate range along y axes.

3 System Calibration

Visual information is the crucial feedback to enable the micro-manipulation and -assembly tasks. Processing visual data determines the path of the end effector in the image frame; however, the input of manipulator is given in its own frame. Thus the mapping between the manipulator frame and the image frame forms a critical component for servoing the probe and the gripper. In order to compute the mapping, a calibration method is developed and implemented.

Several calibration methods exist in the literature that are mostly used in macro scale vision applications [11], [12]. However, these methods cannot directly be employed to calibrate an optical microscope coupled with a CCD camera due to the unique characteristics of the optical system. Large numerical apertures, high optical magnifications and very small depth-of-field property of optical microscopes restricts the calibration to a single parallel plane. Although some methods [13],[14] were proposed to calibrate the optical microscope, they are computationally complex and cannot propose a solution for the close focus microscope, side view, since it does not have the same image forming components with a typical microscope. Therefore a simple modification to Tsai's algorithm [11] is done via weak perspective camera model assumption.

The algorithm used is as follows:

1. The end effector is detected using a template matching method and is started to be tracked.
2. The controller generates number of points for the end effector which corresponds to the corners of a virtual calibration grid.
3. The pixel coordinate of the probe/gripper is computed using the Normalized Cross Correlation (NCC) technique for each given position.
4. Once the positions of the end effector in camera and manipulator space are recorded, the Radial Alignment Constraint (RAC) [11] is employed to compute the rotation and translation from the manipulator coordinate frame to the image coordinate frame.
5. The total magnification (M) of the system and the radial distortion coefficient (κ_1) can be obtained by a least square solution which minimizes the following error:

$$error = \frac{\Sigma^N[(x_i - \tilde{x}_i)^2 + (y_i - \tilde{y}_i)^2]^{1/2}}{N} \tag{1}$$

where N is the total number of points, (x_i, y_i) are image coordinates of a point and $(\tilde{x}_i,\tilde{y}_i)$ are the projected world coordinates for the computed rotation and translation. The average error ranges from 0.01 to 0.08 pixel for the coarse, fine and side views. Since the magnifications of the lenses and cell sizes of cameras can differ, the average error in metric coordinates varies from ten nanometers to less than a micron.

4 Micromanipulation Using Vision and Force

In the literature, several micromanipulation works are presented with no sensor [15], only visual feedback [8],[9] and both vision and force information [1],[7],[6]. Since manipulating an object needs the ability to observe, position, and physically transform the object, both vision and force are essential feedback types for our versatile micro-assembly system. Because the workstation is designed for various microscale applications including biological manipulation, both vision and force control are strongly demanded to ensure successful operation and prevent damage to the objects.

4.1 Visual Feedback

Visual servo control relies on techniques from image processing, estimation and control theory. In our system, several image processing algorithms are developed to detect and track the micro-objects, probe and gripper. Since micro polystyrene balls are used in the experiments, detection of the microspheres at video rate from all the top and side cameras becomes crucial. Having a priori knowledge on the size and shape of the microparticles, the generalized Hough transform which is robust under noise is implemented to detect the spheres, after the connected

components are extracted in the entire frame. Using the procedure, the positions of microballs can be detected in the full frame (640x480 pixel) on the sample glass at a rate of 30 Hz.

It is also crucial to detect and update position of the end effector from the top and side cameras during the experiments. Since the end effector is moved in the 3D, two algorithms are required to detect the position of the probe in the x-z and x-y plane. For the images acquired from the side camera, the probe is detected by using template matching with the NCC comparison method. Because determining the contact point of the probe in the x-y plane is vital in order to locate microspheres precisely, the detection of probe tip is computed in subpixel accuracy by exploiting the known CAD of the probe.

Real-time measurement of the features on the image plane must be done in an efficient, accurate and robust manner. In order to satisfy those criteria, the Kalman filter is employed to estimate the positions of image features given the measurements from the feature extraction module. During the micromanipulation experiments, the locations of two image features, position of the probe tip and the particle which is being pushed are tracked by the tracking module at a frame rate of 30 Hz.

A simple proportional control [16] to ensure an exponential decoupled decrease of the error ($\dot{e} = -\lambda e$) is designed:

$$v = -\lambda {L_s}^{\dagger} e = -\lambda {L_s}^{\dagger}(s - s^*) \tag{2}$$

where and ${L_s}^{\dagger} \epsilon R^{6\times k}$ is the pseudo inverse of interaction matrix L_s, s and s^* are visual features being tracked and desired features respectively, and λ is a positive scalar. It is also possible to design control laws that optimize various system performance measures. An optimal visual control design that aims to minimize the tracking error and the control signal energy can be found in [17].

4.2 Force Feedback

Visual information without force feedback is not adequate for sophisticated micromanipulation tasks which require a high degree of dexterity. Using only visual data can model and control positioning of an object. However, pure position control for delicate or fragile objects such as biological material cannot ensure safe and successful manipulation strategies. Furthermore, force feedback in the micromanipulation tasks can also be used to detect the interaction forces between the end effector and the object [4],[5]. In our system, force from the probe sensor is monitored to continuously push the objects without losing the contact between the probe and the objects during the manipulation tasks. An example plot which demonstrates the interaction force between the microsphere and the force sensing probe for a pulse positioning input is given in Fig.2.

4.3 Path Planning for Collision-Free Transportation of Microobjects

In the previous sections, the required modules to dexterously manipulate the microparticles were presented. Using the feature extraction, tracking, visual and

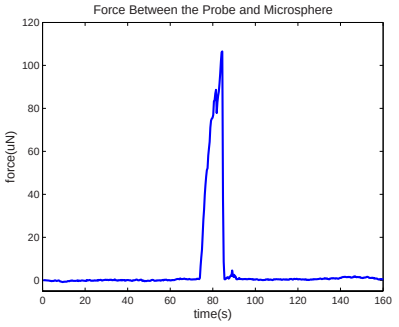
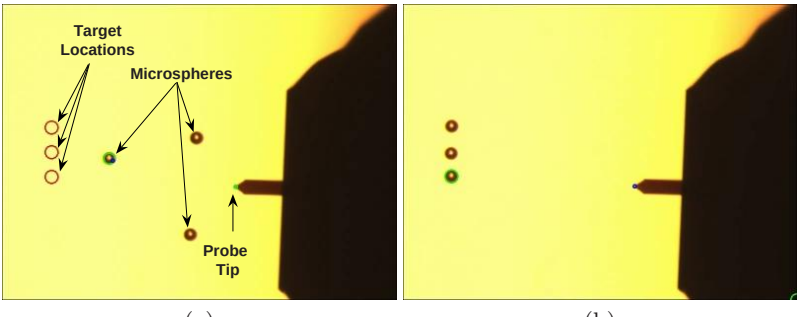


Fig. 2. The Contact Force Between the Probe and the Microsphere

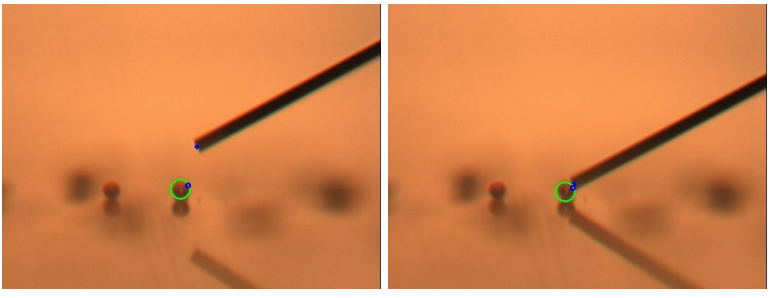


(a) (b)

Fig. 3. Automatic micromanipulation of microspheres, (a) initial configuration of workspace, (b) constructed line pattern

force control modules, we are able to move several microparticles serially along the computed path and generate a pattern by locating them accurately. An illustrative scene for the pattern generation task is given in Fig. 3.

The serial micromanipulation process includes five stages: z movement, explore, obstacle avoidance, pushing and home modes. In the z movement mode, the probe tip is positioned along the z axes by processing the workspace image acquired from the side camera, depicted in Fig. 4.In the explore mode, probe is positioned such that all the workspace is visible and no occlusion occurs due to the probe position. Thus the global particle map can be extracted and used in the path planning. As the probe approaches to the determined particle, there may be an obstacle on the way to the destination. To avoid collision, a rectangular region twice of the size of the obstacle is generated, and the rectangular path is followed by the probe as long as no obstacle exists on the way to the particle. An illustrative figure is shown in Fig. 5. In the home mode, the probe is moved to its initial position with the obstacle avoidance support so that the particle map in the workspace can be updated with no occlusion of the probe.



(a) (b)

Fig. 4. Visual servoing in the x-z plane

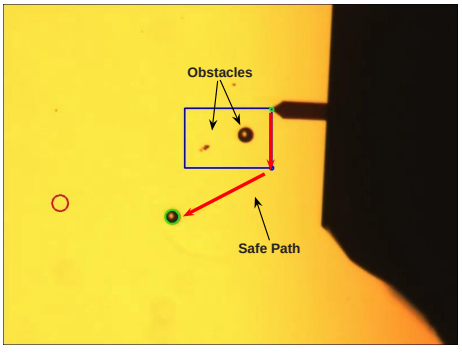


Fig. 5. Obstacle Avoidance

As a result, a micromanipulation experiment in the developed system with suitable hardware and software modules is successfully demonstrated. In the experiment, three microspheres are transported to the desired locations in 45 seconds with 0.7 micron accuracy which is the resolution of the optical microscope.

5 Conclusion

A microassembly workstation which is capable of manipulating the microobjects dexterously by using vision and force is presented. In the vision subsystem, an optical microscope and a lateral view close focus microscope are employed to servo the end effector in 3D by tracking the end effector and microobjects during the manipulation tasks. A force sensing sensor is used to monitor the interaction forces between the probe and the microparticles to keep the pushing force in the optimal range.

Acknowledgments. This work has been supported by SU Internal Grant (IACF06 – 00417). The first author would also like to thank for the support provided by TUBITAK.

References

1. Kim, B., Kang, H., Kim, D.H., Park, G.T., Park, J.O.: Flexible Microassembly System based on Hybrid Manipulation Scheme. In: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Nevada, pp. 2061–2066 (2003)
2. Popa, D., Kang, B.H., Sin, J., Zou, I.: Reconfigurable Microassembly System For Photonics Applications. In: Int. Conf. on Robotics and Automation, Washington, pp. 1495–1500 (2002)
3. Yang, G., Gaines, J.A., Nelson, B.J.: A Flexible Experimental Workcell For Efficient And Reliable Wafer-Level 3D Micro-assembly. In: IEEE Int. Conf. Robotics and Automation, vol. 1, pp. 133–138 (2001)
4. Wang, W.H., Liu, X.Y., Sun, Y.: Autonomous Zebrafish Embryo Injection Using a Microrobotic System. In: Int. Conf. on Automation Science and Engineering, pp. 363–368. IEEE Press, Los Alamitos (2007)
5. Kim, D.-H., Hwang, C.N., Sun, Y., Lee, S.H., Kim, B., Nelson, B.J.: Mechanical Analysis of Chorion Softening in Prehatching Stages of Zebrafish Embryos. IEEE Trans. n Nanobioscience 5(2), 89–94 (2006)
6. Sun, Y., Nelson, B.J.: Microrobotic Cell Injection. In: Int. Conf. on Robotics and Automation, vol. 1, pp. 620–625 (2001)
7. Dionnet, F., Haliyo, D.S., Regnier, S.: Autonomous Micromanipulation using a New Strategy of Accurate Release by Rolling. In: IEEE Int. Conf. on Robotics and Automation, vol. 5, pp. 5019–5024 (2004)
8. Pawashe, C., Sitti, M.: Two-Dimensional Vision-Based Autonomous Microparticle Assembly using Nanoprobes. Journal of Micromechatronics 3(3-5), 285–306 (2006)
9. Ren, L., Wang, L., Mills, J.K., Sun, D.: 3-D Automatic Microassembly by Vision-Based Control. In: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, San Diego, pp. 297–302 (2007)
10. Arai, F., Kawaji, A., Sugiyama, T., Onomura, Y., Ogawa, M., Fukuda, T., Iwata, H., Itoigawa, K.: 3D Micromanipulation System Under Microscope. In: Int. Symposium on Micromechatronics and Human Science, pp. 127–134 (1998)
11. Tsai, R.Y.: A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. IEEE Journal of Robotics and Automation 3, 323–344 (1987)
12. Zhang, Z.Y.: Flexible Camera Calibration by Viewing a Plane from Unknown Orientations. In: IEEE Int. Conf. on Computer Vision, pp. 666–673 (1999)
13. Zhou, Y., Nelson, B.J.: Calibration of A Parametric Model of An Optical Microscope. Optical Engineering 38, 1989–1995 (1999)
14. Ammi, M., Fremont, V., Ferreira, A.: Flexible Microscope Calibration using Virtual Pattern for 3-D Tele micromanipulation. IEEE Trans. on Robotics and Automation, 3888–3893 (2005)
15. Bohringer, K.-F., Donald, B.R., Mihailovich, R., MacDonald, N.C.: Sensorless Manipulation Using Massively Parallel Microfabricated Actuator Arrays. In: IEEE Int. Conf. on Robotics and Automation, pp. 826–833 (1998)

16. Espiau, B., Chaumette, F., Rives, P.: A New Approach to Visual Servoing in Robotics. IEEE Transactions on Robotics and Automation 8(3), 313–326 (1992)
17. Bilen, H., Hocaoglu, M., Ozgur, E., Unel, M., Sabanovic, A.: A Comparative Study of Conventional Visual Servoing Schemes in Microsystem Applications. In: IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, San Diego, pp. 1308–1313 (2007)

HK Segmentation of 3D Micro-structures Reconstructed from Focus

Muhammet A. Hocaoglu and Mustafa Unel

Faculty of Engineering and Natural Sciences, Sabanci University,
34956 Istanbul, Turkey
muhammet@su.sabanciuniv.edu
munel@sabanciuniv.edu

Abstract. This paper presents an evaluation of HK segmentation on 3D micro-structures reconstructed from focus. Due to the necessity of 3D shape and surface structure information for a precise micromanipulation, shape of the object is recovered using shape from focus (SFF). In the SFF procedure, using an image sequence, composed of images captured at different focusing levels, focused image of the micro-structure and its shape are acquired. Then the resulting shape, also called range image, is used in HK segmentation method to extract surface curvature information. Experimental results show that this technique works for both synthetic and real data.

Keywords: HK Segmentation, 3D Micro-structures, Image.

1 Introduction

Precise micromanipulation requires 3D shape and surface structure of the manipulated object. Shape information is extracted from images acquired under the optical microscope. In the literature, there are several active and passive methods like structured light, binocular stereo, shape from shading and shape from focus/defocus to extract 3D geometry of the examined object [1]-[5]. The acquisition of the 3D geometry of micro object enables us to apply a curvature-based descriptor to obtain range surface segmentation [6]-[9]. Curvature-based desriptors were used in several applications. Trucco and Fisher evaluated range image segmentation system using estimates of the sign of the mean and Gaussian curvatures (H and K respectively) on a range image of an automobile part [7]. Cantzler and Fisher compared HK and SC (shape index and curvedness) segmentation on range images of several objects, and it resulted that both methods have the same performance [9]. Moreno et.al applied HK segmentation on 3D human face surfaces for face recognition [10]. These papers addressed the problem of extracting surface information from range images of macro scale objects.

In this paper, shape of the objects were recovered using shape from focus approach [1]. The application of this method requires to move an unknown object on a micro-translation stage with respect to the optical microscope in the z-axis. During this process, at each focus level different parts of the object come into focus

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1173–1180, 2008.
© Springer-Verlag Berlin Heidelberg 2008

and give sharp images. A focus measure operator is used to measure the degree of focus at each pixel of images in the image sequence. Depth estimate for each pixel corresponding to an object point is acquired by interpolating a Gaussian curve to the points around the peak value. Using the resulting shape from this process, curvature-based segmentation of the range image is realized. This segmentation enables to learn whether a surface patch of the range image is planar, cylindric, elliptic or hyperbolic. Having acquired a priori shape and curvature information of the micro object will be helpful in 3D micromanipulation.

The paper is organized as follows: Section 2 defines shape from focus and HK segmentation of range images. Section 3 presents experimental results. Finally, Section 4 concludes the paper with some remarks.

2 3D Reconstruction from Focus and HK Segmentation of Range Images

Curvature-based segmentation methods are used in object classification, pose estimation, and computer graphics in macro world applications. Input to the surface-based descriptor in such applications are range images acquired from the laser scanners. In micro domain although it is possible to apply an active sensing method like 3D Laser Scanning Microscope to have range image of the object, due to its cost, a passive sensing method like shape from focus is more feasible. So, initially a range image is acquired from shape from focus and then a curvature-based segmentation method is applied on the image to extract curvature information of the surface patches.

2.1 Shape from Focus

SFF [1] requires several images of the object at different focusing levels to acquire a 3D map of the object. The object is placed on a micro-translational stage under the optical microscope and the stage is translated in the z-direction until images related to entire 3D object becomes de-focused. In this process, the images are captured at discrete focusing levels with a step size of Δd.

Assuming microscope objective as a thin lens, an object point at a distance d_o from lens will form an image at d_i. If the sensor plane is at a distance d_s and $d_s \neq d_i$, then the image will be a circular blob of radius r_b instead of a point. The relation between r_b and the sensor displacement $\lambda = d_s$ - d_i can be given as $r_b = R\lambda/d_i$ where R is the radius of the lens. Using the Gaussian lens law, the blur radius can be reformulated as [2],

$$r_b = Rd_s[\frac{1}{w_d} - \frac{1}{d_o}] \tag{1}$$

where w_d is working distance of the microscope objective. If an object point is at a distance d_o which is equal to w_d then it forms a focused image. The shape of the object can be estimated by recording the stage position which brings a part of the object into focus.

In order to evaluate the degree of focus, a focus measure has to be selected. The focus measure is calculated in a region around each pixel in the image sequence and the operator gives the highest value at the image which object point is in maximum focus. Sum-modified-Laplacian operator is chosen to be used in the experiments since it performs very well for surfaces that produce textured images [3]. The discrete approximation to the modified laplacian (ML) operator is given by

$$ML(u,v) = |2I(u,v) - I(u-step,v) - I(u+step,v)| + |2I(u,v) - I(u,v-step) - I(u,v+step)| \quad (2)$$

where step is the variable spacing between the pixels. The SML focus measure at a point (i,j) is computed as the sum of modified Laplacian values, in a small window around (i,j), using only values that are greater than a threshold:

$$F(i,j) = \sum_{u=i-W}^{i+W} \sum_{v=j-W}^{j+W} ML(u,v), \quad if \ ML(u,v) > T \quad (3)$$

where W determines the window size around the pixel (i,j) to compute the focus measure and T is the threshold value. A small window size of 3×3 or 5×5 is used since large window size introduces blur by taking into account more pixel values which might be quite different from the pixel in consideration as mentioned in [4]. Computation of $F(i,j)$ at each pixel in the image sequence forms a focus measure profile from which shape can be estimated using Gaussian interpolation.

2.2 HK Segmentation of Range Images

The result of shape from focus algorithm is in the form of a range image which has the same dimensions with the image of the object. To find shape composition of the visible surface of an object, a curvature-based segmentation has to be applied. HK segmentation, which partitions a range image into regions of homogeneous shape, is a proper method for this classification [6].

This method is based on differential geometry, and it uses the mean curvature H and the Gaussian curvature K. Local surface shape can be estimated using signs of H and K at each pixel as shown in Table 1. Calculation of H and K values in a range image S is done according to the following formulas

$$H = \frac{(1+S_x^2)S_{yy} - 2S_xS_yS_{xy} + (1+S_y^2)S_{xx}}{2(1+S_x^2+S_y^2)^{3/2}} \quad (4)$$

$$K = \frac{S_{xx}S_{yy} - S_{xy}^2}{(1+S_x^2+S_y^2)^2} \quad (5)$$

where S_x, S_y, S_{xy}, S_{xx} and S_{yy} are range image derivatives computed from the convolution of numerical derivative filter with the range image. Before these computations, if there is noise in the image due to quantization or resolution,

Table 1. Shape Classification Scheme

H	K	Local Surface Shape
-/0/+	-	hyperbolic
-	0	convex cylindrical
+	0	concave cylindrical
-	+	convex elliptic
+	+	concave elliptic
0	0	plane

gaussian smoothing has to be applied. In this step, as standard deviation of gaussian filter increases, segmentation estimates get worse [7]. A zero-threshold has to be determined for H and K values since numerical estimates of H and K will not be exactly zero [6]. After the computation of H and K images, each pixel is assigned a shape label according to Table 1.

3 Experimental Results

In this section, first the application of SFF algorithm on synthetic and real data is presented, and then surface shape classification results of the range images, acquired from the application of SFF algorithm, using HK segmentation is given.

3.1 SFF Results

SFF is first evaluated on a synthetic texture image shown in Fig. 1(a). An image sequence is created in which at each focus level a different part of the parabolic object is focused. In the creation of these images, camera defocus PSF was modeled as a 2-D Gaussian function and given by $(1)/(2\pi\sigma^2)exp(-(i^2 + j^2)/(2\sigma^2))$ where σ is the deviation of the Gaussian function [2].

SFF algorithm is applied on the image sequence, some of which shown in Fig. 1(b),(c), corresponding to a parabolic shape created assuming $\Delta d = 25\mu m$ and object height is $500\mu m$. The resulting focused image is shown in Fig. 1(d).

Since the spikes in range image would result in problems at HK segmentation stage, these were detected and replaced with the response of median filtering as it is done in [5]. Resulting range image is shown in Fig. 2.

After the application of SFF on synthetic data,it is also evaluated on real data. The image sequence shown in Fig. 3(a)-(c), which belongs to a solder ball on paper, was captured under an optical microscope. Initially as it is shown in Fig. 3(a), paper in the background is focused. Then, micro-translational stage is moved with a $\Delta d = 50\mu m$ until whole object is scanned (Fig. 3(b), (c)). The implementation of SFF on this image sequence resulted in the following focused image shown in Fig. 3(d). Error correction is also applied to the range image acquired from real data. However, due to the rough surface of the ball, the range image result is not very smooth (Fig. 4). Since the object is observed from above, only half of it can be examined. This is why lower part of 3-D map of the object is a cylinder and the top part is spherical.

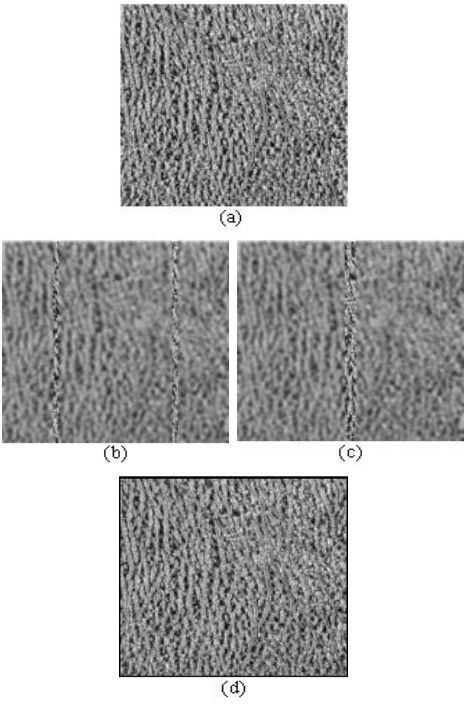


Fig. 1. (a) Original Textured Image (b), (c): Defocused Images in the Image Sequence for SFF (d) Resulting Focused Image from SFF

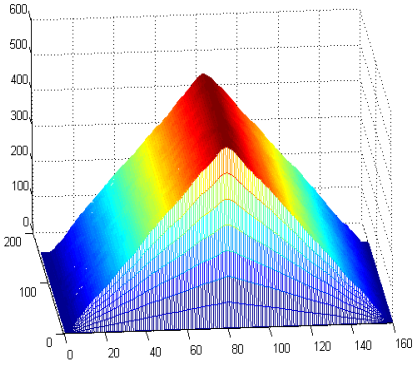


Fig. 2. Resulting Parabolic Range Image from Synthetic Data

3.2 HK Segmentation Results

HK segmentation initially applied on the parabolic profile resulted from the application of SFF on synthetic data to examine the surface structure of the

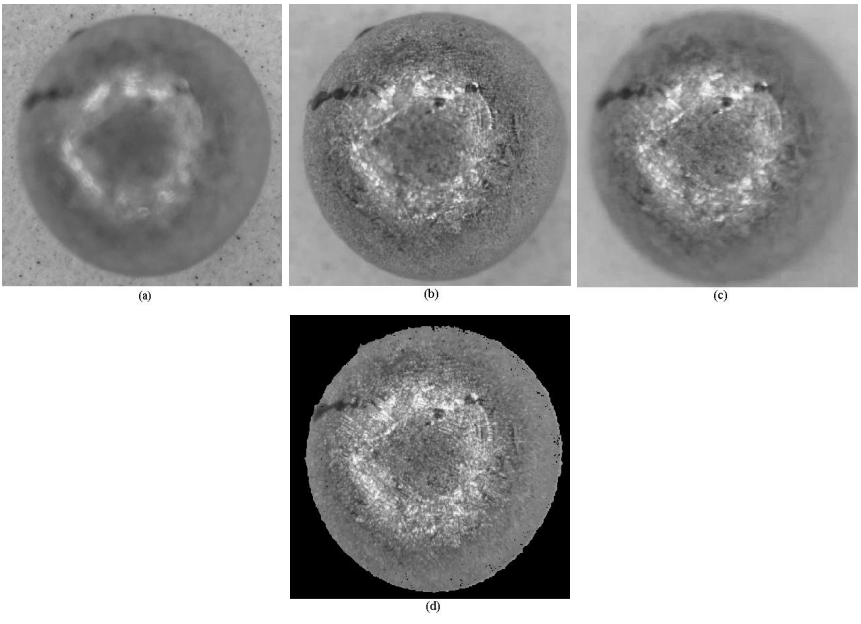


Fig. 3. (a)-(c) Defocused Images of Solder Ball. (d) Resulting Focused Image from SFF.

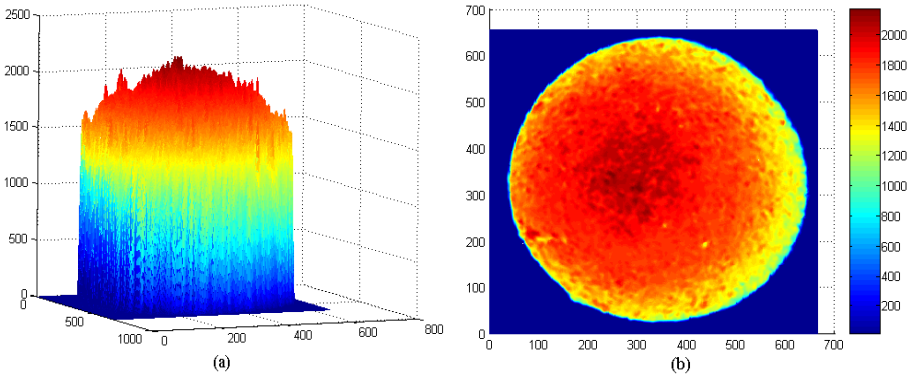


Fig. 4. (a) Solder Ball Range Image (b) Range Image x-y View

parabolic profile. Before the segmentation procedure, gaussian filter with a standard deviation of 1.5 is applied to the range image to smooth the image. Also zero threshold for H and K were determined as 0.03 and 0.0008, respectively. In the resulting segmentation, as it is shown in Fig. 5, left and right of the maximum of the parabola is planar. Although it is expected that the maximum of the surface would be mostly convex cylindric, there are some hyperbolic surface

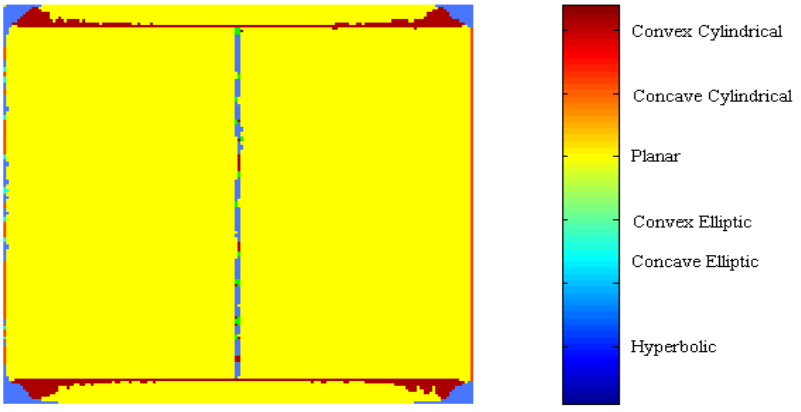


Fig. 5. HK Segmentation of Range Image Result of Synthetic Data

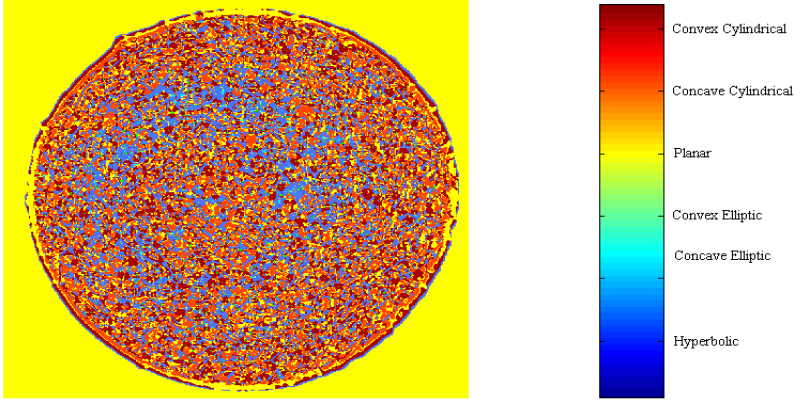


Fig. 6. HK Segmentation of Range Image Result of Real Data

patches detected by HK segmentation. This is due to the errors generated by SFF procedure. Also the edges of the surface are detected as convex cylindrical due to the fact that Gaussian smoothing of the range image results in a convex cylindrical shape at the intersection of parabolic object and background.

Gaussian smoothing was also applied to the SFF result and a zero threshold for H and K was determined. Then surface structure of the solder ball was acquired by applying HK segmentation on the range image result of SFF as it is shown in Fig. 6. Due to the roughness of the surface, mostly concave cylindric patches were detected on the upper surface of the object. There were also some hyperbolic patches again as a result of rough surface structure. However, if the range image was an image of a smooth sphere, it would be convex elliptic.

4 Conclusion

We have shown that HK segmentation can be used on the 3D maps obtained from Shape from Focus (SFF) method to extract surface curvature information. Experimental results show that HK segmentation technique, which was used for surface shape classification of range images of macro-scale objects captured using laser range scanners, can also be used to classify surface curvature and hence related micro-structures. These procedures will be important for classifying/recognizing 3D micro-structures to be manipulated.

Acknowledgments. This work is supported by SU Internal Grant No. IACF06-00417. First author would like to thank for the support provided by TUBITAK.

References

1. Nayar, S.K., Nakagawa, Y.: Shape from Focus. IEEE Transactions on Pattern Analysis and Machine Intelligence 16(8), 824–831 (1994)
2. Pradeep, K.S., Rajagopalan, A.N.: Improving Shape from Focus Using Defocus Cue. IEEE Transactions on Image Processing 16(7), 1920–1925 (2007)
3. Nayar, S.K., Nakagawa, Y.: Shape from focus: An Effective Approach for Rough Surfaces. In: Proceedings of International Conference on Robotics Automation, pp. 218–225 (1990)
4. Malik, A.S., Choi, T.S.: Consideration of Illumination Effects and Optimization of Window Size for Accurate Calculation of Depth Map for 3D Shape Recovery. Pattern Recognition 40, 154–170 (2007)
5. Helmli, F.S., Scherer, S.: Adaptive Shape from Focus with an Error Estimation in Light Microscopy. In: Second International Symposium on Image and Signal Processing and Analysis, pp. 188–193 (2001)
6. Trucco, E., Verri, A.: Introductory Techniques for 3-D Computer Vision. Prentice Hall, New Jersey (1998)
7. Trucco, E., Fisher, R.B.: Experiments in Curvature-Based Segmentation of Range Data. IEEE Transactions on Pattern Analysis and Machine Intelligence 17(2), 177–182 (1995)
8. Besl, P.J., Jain, R.C.: Segmentation Through Variable-Order Surface Fitting. IEEE Transactions on Pattern Analysis and Machine Intelligence 10(2), 167–192 (1988)
9. Cantzler, H., Fisher, R.B.: Comparison of HK and SC Curvature Description Methods. In: Proceedings of Third International Conference on 3-D Digital Imaging and Modeling, pp. 285–291 (2001)
10. Moreno, A.B., Sanchez, A., Velez, J.F., Diaz, F.J.: Face Recognition using 3D Surface-Extracted Descriptors. In: Proceedings of Irish Machine Vision and Image Processing Conference (2003)

Robust Face Tracking Using Bilateral Filtering

Yoon-Hyung Lee, Mun-Ho Jeong, Joong-Jae Lee, and Bum-Jae You

Center for Cognitive Robotics Research

Abstract. This paper introduces a method for face tracking in a video sequence in real time. In this method the profile of color distribution characterizes target's feature. It is invariant for rotation and scale changes. It's also robust to non-rigidity and partial occlusion of the target. We employ the mean-shift algorithm to track the target face and to reduce the computational cost. However, face tracking using color distribution is failed by noises as occlusion including some objects with similar color distribution and with exactly difference color distribution. Thus failures are critical problems. To solve these problems, we employ a bilateral filter which uses the color and range information. We have applied the proposed bilateral filter to track the real time face tracking. The experimental results demonstrate the efficiency of this algorithm. Its performance has been proven superior to the original mean shift tracking algorithm.

Keywords: mean shift, face tracking, occlusion, bilateral filtering.

1 Introduction

The face tracking is an essential field in Human Robot Interaction. The tracker should be robust to partial occlusion, background clutter, changes its size and appearance. And, it is critical that the face tracking algorithm should have low computational cost because it requires real-time processing. Additionally, the majority of the system time is often reserved for high-level tasks such as face recognition, direction of face and mouse or eye detection.

Recently, the mean-shift algorithm [5] has been used as one of the popular methods utilizing the color distribution. This method has an advantage for real-time tracking due to its small computational cost. Sometimes, the face tracking has been failed by occlusions of objects with similar color distributions such as hands and occlusions of objects with exactly different color distributions. The tracking failure, also, can be occurred in the background clutter with similar color distributions. It is challenging work to segment a occluded region in images. To classify the occlusion, there is template matching algorithm [2, 8]. The template matching method yields robustness to searching the occluded pixels. The algorithm, however, is no effect on face tracking with changes of size, appearance and pose, because of matching with a fixed template acquired from the first frame. One of such the methods of estimating and tracking the target position is the particle-filter [4, 9]. In [4], good performances were reported even for the case where the target region in subsequent frames has partially different color distributions. However, these methods require much larger computational costs. So, [6] combined both [4, 9] algorithms and proposed a tracking method which copes

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1181–1189, 2008.
© Springer-Verlag Berlin Heidelberg 2008

with a temporal occlusion. They have a limitation which is temporal occlusion with a small computational cost. However, to classify the occlusion, we did not use a particle filter, because it requires extended computational cost. The color distribution and range information is used for considering the occlusion. Actually, such researches considering the occlusion are few.

In this paper, we combine both bilateral filter and mean-shift algorithm for a real time tracking method which copes with the mentioned occlusions and background clutter. In fact, a bilateral filter [3] is an edge-preserving filter useful for imaging. However, we has been employed its concept.

We first survey the construction of the color distribution model and its similarity measure for the face tracking used in this paper. Then we propose a method of face tracking incorporating the bilateral filter into mean-shift algorithm of regional color distribution. Experimental results show the effective and robust real time tracking of face.

2 Mean-Shift Tracking

In this section, we will first briefly review the concepts behind Mean Shift algorithm. Since the method that we propose in section3 is an extension from our earlier work in [5], we will then provide a quick review of that work.

2.1 Kernel-Based Target Modeling

The color information will be more unstable at peripheral points in the region than its center by being suffered from partial occlusions and background changes. To avoid this instability, many researcher take larger weights for the central points of the region by applying the some kernel. The object model is represented by using kernel histogram of an ellipse region which includes the object.

Let $\{x_i\}_{i=1,\ldots,n}$ be the pixel locations of face image in a frame, centered at position y. In the case of gray level image, the object kernel histogram is expressed by

$$\hat{p}_u(\mathbf{y}) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{\mathbf{y}-\mathbf{x}_i}{h}\right\|^2\right)\delta\left[b(\mathbf{x}_i)-u\right] \quad , \quad u=1,\ldots,m \tag{1}$$

where, $k(\bullet)$ is a profile of Epanechnikov kernel. $\hat{p}_u(y)$ is the value of bin in the color distribution with index u and m is the number of bins. And n_h is region of face. The set $\{\hat{p}_u(y)\}_{u=1,\ldots,m}$ represents the object model. $b(x_i)$ is the quantization value of the pixel intensity at x_i. h is the kernel bandwidth which normalizes coordinates of image so as to make radius of the kernel profile to be one. The constant C_h is derived by imposing the condition $\hat{p}_u(y)\sum_{u=1}^{m} p_u = 1$. In our experiments, the histograms are calculated in the RGB space. To make the low computational complexity, we use 256 bins which determined the upper 3-bits in red space, the upper 3-bits in green space and upper 2-bits in blue space. The kernel histogram $p(y)$ of a target of the tracking object region at y is given with the same manner as q.

Then, we need to evaluate how much does two distribution q and $p(y)$ resemble. We employ the Bhattacharyya coefficient ρ for this evaluation:

$$\hat{\rho}(y) \equiv \rho[\hat{p}(y),\ \hat{q}] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}} \tag{2}$$

2.2 Mean-Shift

The mean-shift algorithm is an efficient method to find the position where a function becomes maximum. In [5], the object tracking is achieved by applying it to the Bhattacharyya coefficient to find the object position.

If the candidate histogram $\{\hat{p}_u(y)\}_{u=1,\ldots,m}$ of the object at the position y is not different largely from its initial value $\{\hat{q}_u\}_{u=1,\ldots,m}$, the Bhattacharyya coefficient can be approximated as (4) by the Taylor expansion and (2):

$$\cdot\ \rho[\hat{p}(\mathbf{y}), \hat{q}] \approx \frac{1}{2}\sum_{u=1}^{m}\sqrt{\hat{p}_u(\mathbf{y}_0)\hat{q}_u} + \frac{C_h}{2}\sum_{i=1}^{n_h}\omega_i k\left(\left\|\frac{\mathbf{y}-\mathbf{x}_i}{h}\right\|^2\right) \tag{3}$$

where

$$\omega_i = \sum_{u=1}^{m}\delta[b(\mathbf{x}_i)-u]\sqrt{\frac{\hat{q}_u}{\hat{p}_u(\mathbf{y}_0)}} \tag{4}$$

To maximize the similarity, second term in equation (3) has to be maximized, first term does not depend on y. The second term represents the density estimate computed with kernel profile k at y in the current frame, with the data being weighted by ω_i. Then for the object tracking, we find the maximum of this second term by the next mean-shift algorithm. Then in the current frame, estimated next location y_{i+1} of region is

$$\mathbf{y}_{i+1} = \frac{\sum_{i=1}^{n}\mathbf{x}_i\omega_i g\left(\left\|\frac{\mathbf{y}_i-\mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n}\omega_i g\left(\left\|\frac{\mathbf{y}_i-\mathbf{x}_i}{h}\right\|^2\right)} \tag{5}$$

where $g(x) = -k'(x)$.

These operations are iterated for every image frame in a video sequence to track a face. Since, in this method, the region of interest is only moved toward the direction in which the Bhattacharyya coefficient becomes large, it may be trapped by a local maximum in the cases of occlusions.

3 Robust Face Tracking Using Bilateral Filtering

3.1 Sensitive Model to Noises

In this section, we describe the detailed effect of 1) occlusion by object with similar color distribution, 2) background with similar color distribution and 3) occlusion by

object with exactly different color distribution. These cases bring about tracking failure and miscomputation of the mentioned high-level implementation. Then, we examine it as noises.

First of all, when the tracker predicts next location of particular face by using mean shift method, color distribution offers very important clue. Therefore when the color distribution is determined, if other object member pixels with similar color distribution is exist in face tracking region, it is regarded as face member pixel. Fig 1. presents the case.

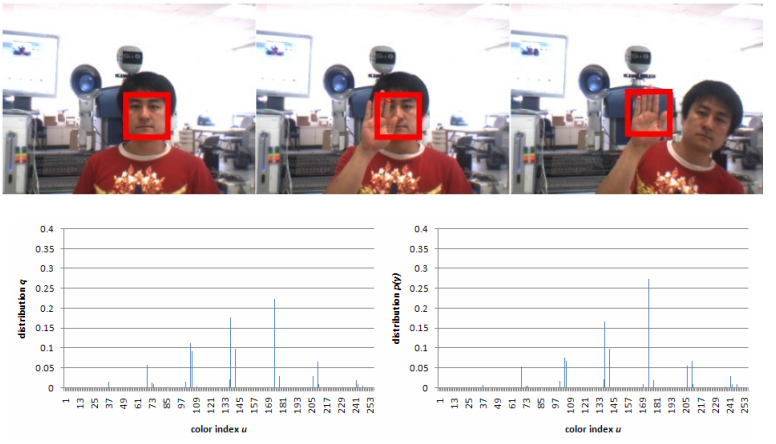


Fig. 1. First row image is a video sequence about 1). The frames 1, 76 and 131 are shown. And second row image is its color distribution.

First row image is an influence of hand with similar color distribution. And second row image is its color distribution at frame 1 and 76, respectively. The similarity between target model q (frame 1) and candidate model p(y) (frame 76) is 0.952. At the frame 131, tracker takes the hand as its target and loses the real target face. Finally, by using the original mean shift algorithm, the face tracking is failed.

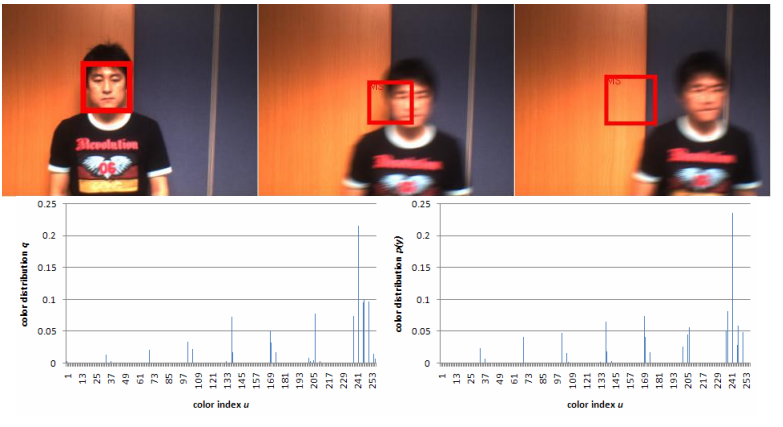


Fig. 2. The first row image is sequence about 2). The frames 1, 211 and 213 are shown. The second row image is its color distribution.

Then, Fig 2. illustrates a tracking result using the original tracker with 2) case which take a background pixels as face member pixels.

First row image is an influence of hand with similar color distribution. And second row image is its color distribution at frame 1 and 211, respectively. The similarity is 0.985 and 0.857, respectively. Finally, by using the original method, the face tracking is failed at frame 213.

Next, there is another type as a noise which some object, with different color distribution, enter to particular face region. Actually this problem is covered by weight function equation (4). When $\omega_i > 1$, the function assign larger weight to *i*th pixel and the estimated location y_{i+1} is moved to new position with the largest similarity. If we track the human body or some object, that is no critical problem. However, in face tracking, accurate face region is required because after track the face, some systems carry out the high-level implementation. (e.g., face recognition, eye detection and prediction of face direction) Since such these tacks depend on face region, accurate face region is very important clue. In fact, 2) occlusion is relative to this reason.

3.2 Bilateral Filtered Support Map

We propose the bilateral filter to solve such problems which has been mentioned in section 3.1.

The face has associated with it a *support map*, which indicates which image pixels are members of a particular face. A range (depth) domain allows easy segmentation of the user from a scene with other people and background objects [1]. We define $s_d(x_i)$ in range domain, the support map for the segmentation, to be

$$s_d(\mathbf{x}_i) = \begin{cases} -1 & \bar{d} - d_i < -3\sigma \\ 1 & \left|\bar{d} - d_i\right| \le 3\sigma \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where, $\bar{d}$ is mean of range in face region which is determined at time *t-1*. And d_i is range value at x_i pixel at current time *t*. An aggregate support map $s_d(x_i)$ over all the above mentioned occlusions is also a useful data structure. Since the support map indicates which image pixels are members of that particular face, background and foreground as occlusion, the aggregate support map represents the segmentation of the image into range classes. Individual pixels are then assigned to particular classes: either to the face class or a background class or the object front of face. A classification decision is made for each pixel by comparing the distance between $\bar{d}$ and d_i, and assigning -1, 1 or 0. We take a Gaussian with variance σ^2.

In the support map, when $s_d(x_i)$ is -1 that means the pixel is occluded by other object with different depth of particular face. And if $s_d(x_i)$ is 1 x_i pixel is the particular face's member. By using support map $s_d(x_i)$, it is useful for segmentation of mentioned occlusion and background in range domain.

However, it is difficult to classify some object with same range information and different color distribution. Therefore, in color domain, we use another support for segmentation of the noise pixels. The equation (7) is to resolve the class membership

likelihood at each pixel into support map $s_c(x_i)$, indicating for each pixel whether individual pixels are face member or not.

$$s_c(\mathbf{x}_i) = \begin{cases} 1 & \omega_i \geq \xi \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

ξ is threshold. A classification decision is made for each pixel by comparing the computed the color distribution q , $p(y)$ and assigning the 1 or 0 at x_i pixel.

Consequently, we denote the bilateral filter which combine the support map $s_d(x_i)$ (in range domain) and the support map $s_c(x_i)$ (in color domain).

$$m(\mathbf{x}_i) = s_d(\mathbf{x}_i) \times s_c(\mathbf{x}_i) \tag{8}$$

Here, $m(x_i)$=1 means that x_i pixel is member pixel of particular face and when $m(x_i)$=0, x_i pixel is occluded by some object. Finally $m(x_i)$=-1 means that x_i pixel is member pixel of background.

3.3 Robust Face Tracking with Bilateral Filtered Support Map

In section 3.1, we introduced a role of weight function ω_i when a tracker predicts next location. We focus on robustly track the particular face by using the bilateral filtered member pixels. To minimize the influence of noises with the exactly different color distribution at x_i, we define a condition as follows:

$$\omega_i' = \begin{cases} \omega_i & : m(\mathbf{x}_i) = 1 \\ 1 & : m(\mathbf{x}_i) = 0 \\ 0 & : m(\mathbf{x}_i) = -1 \end{cases} \tag{9}$$

In equation (9), we can easily know that whether each pixel is occluded or not as. When $m(x_i)$=0 , means that x_i pixel is occluded by some objects. Then if ω_i' is 1, when estimate the next position, we can reduce the effect of x_i as noise. And if $m(x_i)$=-1, x_i is background class member. In this case the pixel is rejected as ω_i'=0, in tracking face. The equation (9) means that weight function is filtered by bilateral filter.

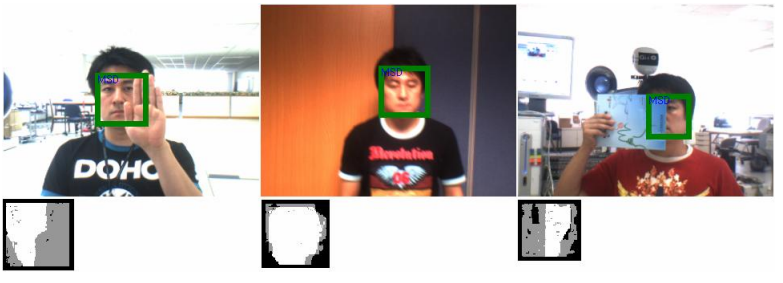


Fig. 3. The performance of redefined ω_i' in the introduced cases 1), 2) and 3)

Fig.3 shows the performance of bilateral filtered weight. In second row image, a white point means ω_i'= ω_i as a particular face member pixel, a gray point means ω_i'=1 as a occluded pixel and black point is ω_i'=0 as a background or miscomputed range data.

Given the above representation (9), we can be rewritten as

$$y_1 = \frac{\sum_{i=1}^{n} \mathbf{x}_i \omega_i' g\left(\left\|\frac{\mathbf{y}_0 - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} \omega_i' g\left(\left\|\frac{\mathbf{y}_0 - \mathbf{x}_i}{h}\right\|^2\right)} \tag{10}$$

Then we recomputed the mean of range $\bar{d}$ at time *t*.

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n_h} d_i \times m(\mathbf{x}_i) \tag{11}$$

where, *n* is the number of pixel which is member pixel.

4 Experiments

In our experiments, a face color distribution computed by Epanechnikov kernel has been derived in the space with upper 3bits in red space, upper 3bits in green space and upper 2bits in blue space. i.e. it has 256bins.

We have been addressed the problems about occlusions in Section 3.1 and implemented the algorithm for 1) the occlusions by the object with similar color distribution such as hands, 2) background clutter with similar color distribution and 3) the occlusions by object with exactly different color distribution and same rage. They were applied for face tracking in appropriate video clips made. For detecting the face, the Adaboost algorithm [7] has been used. When the face detect, target face *q* is initialized.

Fig. 4 presents the result of tracking in Fig. 1 video sequence. The block shows the tracking result using a proposed bilateral filter, and the target face does not lose.

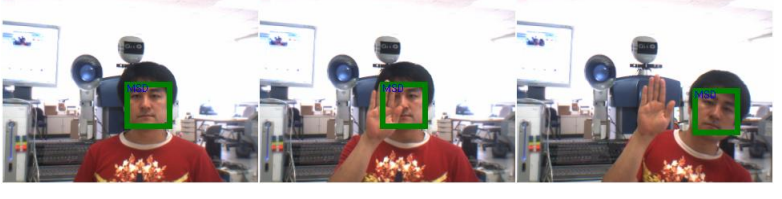


Fig. 4. The image is a video sequence about 1) by using proposed method. The frames 1, 76 and 131 are shown.

And, Fig. 5 shows three frames of face tracking sequences with a background including similar color distribution. We already had shown the tracking result by general mean-shift using candidate model without the bilateral filtering in Fig. 2. When the 2) partial occlusion occurs as we defined, the tracker takes the background as its target and loses the real target face. The sequence shows the tracking result with our proposed method, and the target does not lose the severe occlusion at frame 213.

Fig. 5. The image is sequence about 2). The frames 1, 211 and 213 are shown.

Next example is shown in Fig. 6, demonstrating the performance of tracker, with the bilateral filter, to reduce the influence of noise pixels like as blue book. The book has same rage information and different color distribution. For convenience, before blue book enters to face region, first image is shown. The face region is shifted to location by general mean shift in second image. The third image shows the proposed tracking method. It is kept up the face region. A last image shows the bilateral filtered face member pixels as white point and white point relate with the pixels which is regard to $s_d(x_i)=0$.

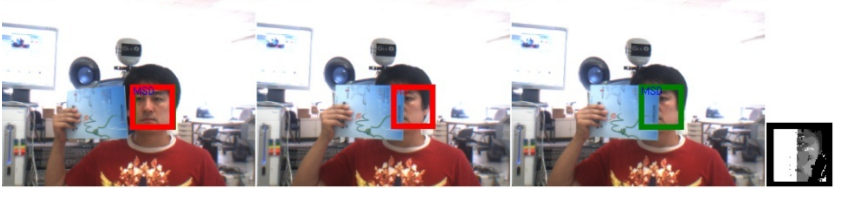


Fig. 6. An object with same rage information and exactly different color distribution. The frames 64, 100 are shown.

The implementation of our tracking method has been simulated on a 3GHz PC where the tracker runs at a rate of 15fps allowing tracking a face. As opposed to our method, general mean-shift algorithm runs at a rate of 0.015 sec and ours runs at a rate of 0.016 sec. Consequently our tracking process satisfies the real-time requirement.

5 Conclusion

This paper proposes a method for mean-shift based real-time face tracking in video sequence with color and range information. The proposed bilateral filter consists of two support map in range domain and color domain. A support map made in range domain is easily classified the particular face, background and occlusion. However, because the support map cannot classify the occlusion by objects with same rage information and exactly different color distribution, we defined other support map in color domain. Finally, the proposed bilateral filter combined two support maps is useful for environment with the occlusions by objects which consist of similar color distribution or exactly different color distribution, with background clutter with similar color distribution. Moreover, the tracker can handle severe occluded pixels in

complexity scenes by the bilateral filtering with the mean-shift tracking algorithm and bilateral filter for estimating the face. Also, the tracking process satisfies the real-time request. When occlusions occur during face tracking, our method can achieve the robustness.

Acknowledgments. This work was supported by the IT R&D program of MIC/IITA [2006-S-028-01, Development of Cooperative Network-based Humanoids Technology].

References

1. Darrell, T., Gordon, G., Harville, M.: Integrated person tracking using stereo, color, and pattern detection. In: CVPR 1998, Santa Barbara, California (1998)
2. Li, B., Chellappa, R.: Simultaneous Tracking and Verification via Sequential Posterior Estimation. Proc. Computer Vision and Pattern Recognition 2, 110–117 (2000)
3. Tomasi, C., Manduchi, R.: Bilateral Filtering for Gray and Color Images. In: Proceedings of the 1998 IEEE International Conference on Computer Vision, Bombay, India (1998)
4. Nummiaro, K.E., Koller-Meier, Van Gool, L.: Color Features for Tracking Non-Rigid Objects. ACTA Automatica Sinica (2003)
5. Dorin, C., et al.: Kernel-Based Object Tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(5) (May 2003)
6. Deguchi, K., Kawanaka, Okatani, O.T.: Object tracking by the mean-shift of regional colour distribution combined with the particle-filter algorithm. In: Proc. ICPR 2004, pp. 506–509 (2004)
7. Freund, Y., Robert, Schapire, E.: A decision-theoretic generalization of on-line learningand an application to boosting. Journal of Computer and System Sciences 55(1), 119–139 (1997)
8. Hager, G.D., Belhumeur, P.N.: Efficient Region Tracking with Parametric Models of Geometry and Illumination. IEEE Trans. Pattern Analysis and Machine Intelligence 20(10), 1025–1039 (1998)
9. Isard, M., Blake, A.: CONDENSATION-Conditional Density Propagation for Visual Tracking. International Journal on Computer Vision 29(1), 5–28 (1998)

Fuzzy Neural Network-Based Sliding Mode Control for Non-spinning Warhead with Moving Mass Actuators

Hongchao Zhao[1], Ruchuan Zhang[2], and Ting Wang[1]

[1] Faculty 703, Naval Aeronautical and Astronautical University, 264001 Yantai, China
navyzhc@163.com, twang703@sina.com
[2] Postgraduate Brigade, Naval Aeronautical and Astronautical University, 264001 Yantai, China
fxqyjs@sina.com

Abstract. Dynamic model of a non-spinning reentry warhead with moving mass actuators is analyzed. The coupling terms between pitch channel and yaw channel are regarded as the uncertainties of each channel, thus the dynamic model is simplified. In order to approximate the bound of lumped uncertainty, two fuzzy neural network-based sliding mode controllers are designed for the moving mass control systems of pitch channel and yaw channel. They control the motion of two moving masses and realize the exact control of warhead's attitude angles. The effect of coupling terms is overcome, and thus the robustness and control precision of the moving mass control systems are improved. Simulation results show the designed fuzzy neural network-based sliding mode controllers have good control effect.

1 Introduction

With the increasing development of ballistic missile defense system, the penetration ability of a ballistic missile is facing with severe challenges. Among all kinds of penetration schemes the terminal maneuver of reentry warhead is of high effect and of low cost. The control equipments used to perform terminal maneuver of reentry warheads include aerodynamic fins, lateral jet motors and moving mass actuators. The moving mass actuators execute the terminal maneuver control task by driving moving masses inside the warhead [1], [2], [3]. Comparing with aerodynamic fins and lateral jet motors, the moving mass actuators have many advantages, so it is paid great attention to recently. However the moving mass control technique is still at the stage of theoretical research, and many theoretical and practical problems are yet unresolved.

The non-spinning warhead utilizes difference fins to stabilize the roll channel. The pitch channel and yaw channel are controlled by two independent masses which move along different lines. Former researchers [4], [5], [6] mainly research the dynamic model of a non-spinning reentry warhead, but their researches do almost not involve the control methods. It is an important problem that which control method should be adopted to control the moving mode of masses. There are many coupling elements between the pitch channel and yaw channel of a reentry warhead. In the course of reentry flight the warhead suffers from much external disturbance. The advanced

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1190–1197, 2008.
© Springer-Verlag Berlin Heidelberg 2008

robust control method must be adopted to exactly control the motion of masses, in order to perform exact control of flight attitude and maneuvering trajectory of a reentry warhead and improve terminal attack precision.

In the recent decades, many scholars have researched the fuzzy neural network-based sliding mode control method [7], [8], [9]. In this method a sliding mode controller is used to improve the stability and speediness of control system, and a fuzzy neural network is used to adjust the parameter of the sliding mode controller in order to compensate the system's uncertainties and to improve its robustness. Because of above advantages we apply this method to design the warhead's moving mass control system. The warhead's attitude is exactly controlled by controlling the motions of the two masses.

2 Model of Non-spinning Warhead

Since the time of non-spinning warhead's reentry flight is short, the impact of the Earth's rotation on the warhead is so little as to be omitted. The roll channel is stabilized by difference fins, hence

$$\gamma = 0,\ \omega_{x1} = 0. \tag{1}$$

The coupling terms between pitch channel and yaw channel are regarded as both uncertainties, and the warhead's dynamic model in body coordinates in [6] can be transformed into the following simple form:

$$\begin{aligned}\bar{J}_{y1}\dot{\omega}_{y1} = {} & M_{y1} + (u_p + u_q)lZ + u_q\delta_z X + (u_p m_q - u_q m_p)l\delta_z\omega_{y1}^2 - \\ & 2(1-u_q)m_q\omega_{y1}\delta_z\dot{\delta}_z + (1-u_p-u_q)m_q l\ddot{\delta}_z + \Delta_1(\omega_{z1}, \delta_y)\end{aligned}, \tag{2}$$

$$\begin{aligned}\bar{J}_{z1}\dot{\omega}_{z1} = {} & M_{z1} + (u_p + u_q)lY - u_q\delta_y X + (u_p m_q - u_q m_p)l\delta_y\omega_{z1}^2 - \\ & 2(1-u_p)m_p\omega_{z1}\delta_y\dot{\delta}_y - (1-u_p-u_q)m_p l\ddot{\delta}_y + \Delta_2(\omega_{y1}, \delta_z)\end{aligned}, \tag{3}$$

where

$$\bar{J}_{y1} = J_{y1} + (1-u_p-u_q)(m_p+m_q)l^2 + (1-u_q)m_q\delta_z^2, \tag{4}$$

$$\begin{aligned}\Delta_1(\omega_{z1}, \delta_y) = {} & (1-u_p-u_q)m_p l\delta_y\omega_{y1}\omega_{z1} - 2u_q m_p\omega_{z1}\delta_z\dot{\delta}_y - \\ & u_q m_p l\delta_z\omega_{z1} + (1-u_q)m_q l\delta_z\omega_{z1}^2\end{aligned}, \tag{5}$$

$$\bar{J}_{z1} = J_{z1} + (1-u_p-u_q)(m_p+m_q)l^2 + (1-u_p)m_p\delta_y^2, \tag{6}$$

$$\begin{aligned}\Delta_2(\omega_{y1}, \delta_z) = {} & [u_p m_q - (1-u_p)]m_p l\delta_y\omega_{y1}^2 + 2u_p m_q\omega_{y1}\delta_y\dot{\delta}_z - \\ & (1-u_p-u_q)m_q l\delta_z\omega_{y1}\omega_{z1}\end{aligned}, \tag{7}$$

where J_{y1} and J_{z1} are y_b-, z_b-components of the moment of inertia in the body coordinates, respectively; ω_y and ω_z are y_b-, z_b-components of angle velocity vector, respectively; M_{y1} and M_{z1} are y_b-, z_b-components of the aerodynamic moment, respectively; X, Y and Z are x_b-, y_b-, z_b-components of the aerodynamic force, respectively; m_p and m_q are masses of the two masses; u_p and u_q are mass ratios of the two masses; the position vectors of masses p and q in the body coordinates are $[l \quad \delta_y \quad 0]^T$ and $[l \quad 0 \quad \delta_z]^T$, respectively; $\Delta_1(\omega_{z1}, \delta_y)$ and $\Delta_2(\omega_{y1}, \delta_z)$ are uncertainties (i.e. coupling terms) of pitch channel and yaw channel, respectively.

From (1) the kinematic model of the warhead's attitude is simply expressed as

$$\dot{\psi} = \omega_{y1}, \tag{8}$$

$$\dot{\varphi} = \omega_{z1} / \cos\psi, \tag{9}$$

where φ and ψ are pitch angle and yaw angle of the warhead, respectively.

The model of pitch channel consists of (3) and (9). Its output variable is pitch angle φ, and its control input is the acceleration $\ddot{\delta}_y$ of mass p. The model of yaw channel consists of (2) and (8). Its output variable is yaw angle ψ, and its control input is the acceleration $\ddot{\delta}_z$ of mass q. Note that $|\psi| < \pi/2$ during a warhead's reentry flight phase, so we have $\cos\psi > 0$.

3 Fuzzy Neural Network-Based Sliding Mode Controller Design

3.1 Sliding Mode Controller Design

We treat the coupling terms as uncertainties of each channel, thus pitch channel and yaw channel are independent of each other. The moving mass control system of each channel can be designed independently. Since the warhead has pitch and yaw symmetry, we take the pitch channel as an example and design its controller to control the acceleration of mass p.

Considering the output tracking problem of pitch angle, we define the tracking error and its derivative as follows,

$$e_\varphi = \varphi - \varphi^*, \ \dot{e}_\varphi = \dot{\varphi} - \dot{\varphi}^*, \tag{10}$$

where φ^* is the command signal of pitch angle. A sliding surface of the sliding mode control is defined as

$$s_1 = c_1 e_\varphi + \dot{e}_\varphi, \ c_1 > 0. \tag{11}$$

The reaching condition of the sliding mode control system is as follows,

$$\dot{V} < 0 \ (s_1 \neq 0), \ V = s_1^2 / 2. \tag{12}$$

Taking the derivative of (11) and substituting it into (12), we get

$$\begin{aligned}\dot{V} &= s_1\dot{s}_1 = s_1[c_1\dot{e}_\varphi + \ddot{e}_\varphi] \\ &= s_1\{c_1\dot{e}_\varphi - \ddot{\varphi}^* + [M_{z1} + (u_p + u_q)lY - u_q\delta_y X + (u_p m_q - u_q m_p)l\delta_y\omega_{z1}^2 - \\ &\quad 2(1-u_p)m_p\omega_{z1}\delta_y\dot{\delta}_y - (1-u_p-u_q)m_p l\ddot{\delta}_y]/(\bar{J}_{z1}\cos\psi) + \bar{\Delta}_2(\omega_{y1}, \delta_z)\}\end{aligned} \tag{13}$$

where

$$\bar{\Delta}_2(\omega_{y1}, \delta_z) = [\Delta_2(\omega_{y1}, \delta_z) + \omega_{y1}\omega_{z1}\tan\psi/\cos\psi]/(\bar{J}_{z1}\cos\psi) \tag{14}$$

is called "lumped uncertainty". It is assumed to be bounded, i.e.

$$|\bar{\Delta}_2(\omega_{y1}, \delta_z)| \le D(t), \tag{15}$$

where $D(t)$ is a continuous function. If $D(t)$ is known beforehand, then the sliding mode controller can be designed as

$$\begin{aligned}\ddot{\delta}_y = \frac{1}{(1-u_p-u_q)m_p l}[&\bar{J}_{z1}\cos\psi(c_1\dot{e}_\varphi - \ddot{\varphi}^*) + M_{z1} + (u_p+u_q)lY - u_q\delta_y X + \\ &(u_p m_q - u_q m_p)l\delta_y\omega_{z1}^2 - 2(1-u_p)m_p\omega_{z1}\delta_y\dot{\delta}_y + \bar{J}_{z1}\cos\psi(D(t)+\rho)\mathrm{sgn}(s_1)]\end{aligned}, \tag{16}$$

where $\rho > 0$, $\mathrm{sgn}(s_1)$ is a sign function. Substituting (16) and (15) into (13) gives

$$\begin{aligned}\dot{V} &= s_1[-(D(t)+\rho)\mathrm{sgn}(s_1) + \bar{\Delta}(\omega_{y1}, \delta_z)] \\ &\le |s_1||\bar{\Delta}(\omega_{y1}, \delta_z)| - D(t)|s_1| - \rho|s_1| \\ &\le -\rho|s_1| < 0 \quad (s_1 \ne 0)\end{aligned} \tag{17}$$

From (17) the reaching condition (12) is guaranteed. However, the bound of lumped uncertainty is difficult to obtain in advance for some variables are unknown. Therefore a fuzzy neural network controller is introduced to estimate the bound. Compared with conventional estimator, the fuzzy neural network controller utilizes prior expert knowledge to accomplish control object more efficiently.

3.2 Fuzzy Neural Network Controller Design

Fuzzy neural network is generally a fuzzy inference system constructed from the structure of neural network. Learning algorithms are used to adjust the weights of the fuzzy inference system [7]. Its output $o(x)$ can be expressed as:

$$o(x) = \frac{\sum_{i=1}^{h}\bar{y}_i\left(\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)\right)}{\sum_{i=1}^{h}\left(\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)\right)} = v^T\phi(x), \tag{18}$$

where $\mu_{A_j^i}(x_j)$ is the membership function of fuzzy variable x_j, h is the total number of the fuzzy rules, $\bar{y}_i$ is the point at which $\mu_{B^i}(\bar{y}_i) = 1$, A_j^i and B^i denote fuzzy sets

of input and output variable, respectively, $v=[\bar{y}_1\ \ \bar{y}_2\ \cdots\ \bar{y}_h]^T$ is a weight vector of the output layer, $x\in\Re^n$ is input vector, and $\phi(x)=[\phi^1\ \ \phi^2\ \cdots\ \phi^h]^T$ is a fuzzy basis vector, where ϕ^i is defined as

$$\phi^i(x)=\frac{\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)}{\sum_{i=1}^{h}\left(\Pi_{j=1}^{n}\mu_{A_j^i}(x_j)\right)}. \tag{19}$$

The fuzzy neural network in the form of (18) can be used as an approximator to approximate the uncertain nonlinear function to arbitrary accuracy as Lemma 1.

Lemma 1 [8]**.** For any given real continuous function χ on a compact set $U\subset\Re^n$ and arbitrary small $\varepsilon>0$, there exists an optimal parameter vector v^*, such that

$$\sup_{x\subset U}|v^{*T}\phi(x)-\chi(x)|<\varepsilon. \tag{20}$$

Now we use the fuzzy neural network to approximate $D(t)$. Its input vector is $x=[e_\varphi\quad \dot{e}_\varphi]^T$, and its output is $o(x)$. Form Lemma 1, for arbitrary small $\varepsilon>0$, there exists an optimal parameter vector v^*, $o^*(x)=v^{*T}\phi(x)$, such that

$$\sup_{x\subset U}|v^{*T}\phi(x)-D(t)|<\varepsilon, \tag{21}$$

which yields

$$D(t)<v^{*T}\phi(x)+\varepsilon. \tag{22}$$

However, v^* can't be used directly, so we introduce $\hat{v}$ as an estimate of v^*. The design of fuzzy neural network-based sliding mode controller and the tracking performance analysis are stated in the following theorem

Theorem 1. For the pitch channel (3) and (9), a sliding surface of the sliding mode control is defined as (11), the fuzzy neural network-based sliding mode controller is designed as

$$\begin{aligned}\ddot{\delta}_y=\frac{1}{(1-u_p-u_q)m_p l}\big[&\bar{J}_{z1}\cos\psi(c_1\dot{e}_\varphi-\ddot{\varphi}^*)+M_{z1}+(u_p+u_q)lY-u_q\delta_y X+\\&(u_p m_q-u_q m_p)l\delta_y\omega_{z1}^2-2(1-u_p)m_p\omega_{z1}\delta_y\dot{\delta}_y+\bar{J}_{z1}\cos\psi(\hat{v}^T\phi(x)+\rho)\mathrm{sgn}(s_1)\big]\end{aligned}, \tag{23}$$

and the learning algorithm of the weight vector is chosen as

$$\dot{\hat{v}}=\Gamma|s_1|\phi(x), \tag{24}$$

where $\Gamma=diag(\eta_1\ \ \eta_2\ \cdots\ \eta_h)$ is the learning rate matrix, then the error e_φ will converge to zero exponentially.

Proof. Choose the following Lyapunov function candidate:

$$V = s_1^2 / 2 + \tilde{v}^T \Gamma^{-1} \tilde{v} / 2 , \tag{25}$$

where $\tilde{v} = \hat{v} - v^*$, and $\dot{\tilde{v}} = \dot{\hat{v}}$. Using (22), (23) and (24), the time derivative of (25) is as follows,

$$\begin{aligned} \dot{V} &= s_1 \dot{s}_1 + + \tilde{v}^T \Gamma^{-1} \dot{\tilde{v}} \\ &= s_1 [\overline{\Delta}_2(\omega_{y1}, \delta_z) - (\hat{v}^T \phi(x) + \rho) \operatorname{sgn}(s_1)] + \tilde{v}^T \Gamma^{-1} \dot{\hat{v}} \\ &\le | s_1 | [v^{*T} \phi(x) + \varepsilon] - | s_1 | [\hat{v}^T \phi(x) + \rho] + \tilde{v}^T \Gamma^{-1} \dot{\hat{v}} \\ &= (\varepsilon - \rho) | s_1 | - \tilde{v}^T \phi(x) | s_1 | + \tilde{v}^T \Gamma^{-1} \dot{\hat{v}} \\ &= (\varepsilon - \rho) | s_1 | < 0 \quad (s_1 \neq 0) \end{aligned} , \tag{26}$$

where $\varepsilon - p < 0$ is certainly satisfied because ε is arbitrary small. Consequently the reaching condition of the sliding mode control system is guaranteed, and the system states can converge to the sliding surface and then stay in it, i.e.

$$s_1 = c_1 e_\varphi + \dot{e}_\varphi = 0 , \; e_\varphi(t) = e_\varphi(0) e^{-c_1 t} . \tag{27}$$

This means the error e_φ will converge to zero exponentially. This completes the proof.

The fuzzy neural network-based sliding mode controller of yaw channel is easily designed according to that of pitch channel. Two moving mass actuators fix in the base of the non-spinning warhead and parallel the y_b-axis and z_b-axis respectively. One drives mass p to move up and down inside the warhead, the other drives mass q to move horizontally inside the warhead. Their effective driving forces are as follows, respectively,

$$F_p = m_p \ddot{\delta}_y , \; F_q = m_q \ddot{\delta}_z . \tag{28}$$

4 Numerical Simulation

In order to evaluate the performance of the fuzzy neural network- based sliding mode controller, a non-spinning warhead is illustrated in the simulation. Its parameters are $M = 500\,\text{kg}$, $L = 1.5\,\text{m}$, and $J_{y1} = J_{z1} = 100\ \text{kg·m}^2$. Two moving mass actuators are same, and both parameters are $m_p = m_q = 30\,\text{kg}$, $l = 0.3\,\text{m}$. The motions of two masses are controlled by two fuzzy neural network-based sliding mode controllers. In the simulation, the warhead's initial attitude angles are assumed as $\varphi_0 = -45^\circ$ and $\psi_0 = 0^\circ$. The commands of the warhead's attitude angles are $\varphi^* = -45^\circ + 5^\circ \sin(\pi t)$ and $\psi^* = 5^\circ \sin(\pi t)$ in the terminal manoeuver phase. Simulation results are shown in Figures 1 and 2.

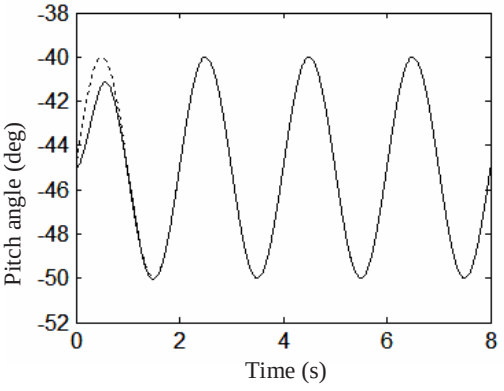


Fig. 1. Command (*dotted line*) and tracking curve (*real line*) of pitch angle

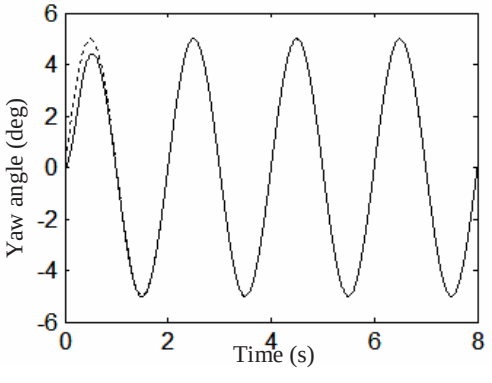


Fig. 2. Command (*dotted line*) and tracking curve (*real line*) of yaw angle

Figure 1 depicts the command and tracking curve of the warhead's pitch angle. Figure 2 depicts the command and tracking curve of its yaw angle. From the simulation results, the pitch angle and yaw angle can track both commands quickly and exactly. That shows the designed fuzzy neural network-based sliding mode controllers can overcome the effect of coupling terms and improve the speediness, robustness and control precision of the moving mass control system.

5 Conclusions

In this paper the non-spinning warhead with moving mass actuators is analyzed. A fuzzy neural network-based sliding mode control method is developed for the design of moving mass control system to control the motions of two masses. This method introduces a fuzzy neural network approximator into the sliding mode controller to approximate the bound of lumped uncertainty, which overcomes the effect of coupled terms and improves the speediness, robustness and control precision of the moving

mass control system. Simulation results show that the designed fuzzy neural network-based sliding mode controller has good control effect.

Acknowledgment

This research is supported partially by China Postdoctoral Science Foundation #20070421095.

References

1. Robinett, R.D., Rainwater, B.A., Kerr, S.A.: Moving Mass Trim Control for Aerospace Vehicles. Journal of Guidance, control, and Dynamics 19, 1064–1070 (1996)
2. Menon, P.K., Sweriduk, G.D., Ohlmeyer, E.J., Malyevac, D.S.: Integrated Guidance and Control of Moving Mass Actuated Kinetic Warheads. Journal of Guidance, control, and Dynamics 27, 118–126 (2004)
3. Wang, S., Yang, M., Wang, Z.: Moving-Mass Trim Control System Design for Spinning Vehicles. In: Proceedings of the 6th World Congress on Intelligent Control and Automation, Dalian, China, pp. 1034–1038 (2006)
4. Whitacre, W.: Rigid Body Control Using Internal Moving Mass Actuator, http://www.aoe.vt.edu/~cwoolsey/Advisees/Undergraduate/WhitacreVSGCPaper.pdf
5. Lin, P., Zhou, F., Zhou, J.: Moving Centroid Control Mode for Reentry Warhead. Aerospace Control 25, 16–20 (2007)
6. Li, R., Jing, W., Gao, C.: Dynamics Modelling and Simulation for Moving-Mass Reentry Vehicle. In: Proceedings of the 26th Chinese Control Conference, Zhangjiajie, China, pp. 497–501 (2007)
7. Wang, W.Y., Cheng, C.Y., Leu, Y.G.: An Online GA-Based Output-feedback Direct Adaptive Fuzzy-Neural Controller for uncertain nonlinear systems. IEEE Transactions on Systems, Man, and Cybernetics 34, 334–345 (2004)
8. Wang, W.Y., Chan, M.L., James Hsu, C.C., et al.: H∞ Tracking-Based Sliding Mode Control for Uncertain Nonlinear Systems via an Adaptive Fuzzy-Neural Approach. IEEE Transactions on Systems, Man, and Cybernetics 32, 483–492 (2002)
9. Zhao, H., Yu, H., Gu, W.: Fuzzy Neural Network-Based Sliding Mode Control for Missile's Overload Control System. In: Proceedings of 2005 International Conference on Neural Network & Brain, Beijing, China, pp. 1786–1790 (2005)

A Novel Multi-robot Coordination Method Based on Reinforcement Learning

Jian Fan[1, 2], MinRui Fei[1], LiKang Shao[3], and Feng Huang[4]

[1] Shanghai Key Laboratory of Power Station Automation Technology, College of Mechatronics Engineering and Automation, Shanghai University , Shanghai 200072, China
[2] Operations Research Centre, Nanjing Army Command College, Nanjing 210045, China
[3] PLA Artillery Academy, HeFei 230031, China
[4] Staff Room of People's Air Defence Command and Communications, PLA Institute of Communication and Command,Wuhan 430010, china
jfan@mail.shu.edu.cn, mrfei@staff.shu.edu.cn,
shaolikang@163.com, hf0791@126.com

Abstract. Focusing on multi-robot coordination, role transformation and reinforcement learning method are combined in this paper. Under centralize control framework, the distance nearest rule which means that the nearest robot ranges from obstacles is selected to be the master robot for controlling salve robots is presented. Meanwhile, different from traditional way which reinforcement learning is applied in online learning of multi-robot coordination, this paper proposed a novel behavior weight method based on reinforcement learning, the robot behavior weights are optimized through interacting with environment and the coordination policy based on maximum behavior value is presented to plan the collision avoidance behavior of robot. The learning method proposed in this paper is applied to the application related to collaboration movement of mobile robots and demonstrated by the simulation results presented in this paper.

Keywords: reinforcement learning, behavior weight, role transformation, multi-robot.

1 Introduction

Since 1990s, coordination and cooperation of multi-robot have been the inevitable technology development trend in robot research areas; many scientific organizations spread researches in this area. Parker [1] proposed a distributed behavior framework (L-ALLIANCE) in multi-robot coordination. "Target terminating" concept and master-slave control framework is imported in [2]. Huntsberger[3] developed a multi-robot framework based on behavior "CAMPOUT" and used it in NASA robot project. Yamada [4] designed a behavior selecting method based on rules. Under unknown static environment, Miyata [5] also developed a dynamic risk plan framework. But all these methods have a common disadvantage: robots have no learning ability and can not improve their ability through experience. Normally, a robot with fixed ability can not adopt the environment under unknown dynamic environment.

The concept of reinforcement learning is firstly proposed in 1950s [6], it's be the important research work of robot until 1980s. With the development of multi-robot

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1198–1205, 2008.
© Springer-Verlag Berlin Heidelberg 2008

application, reinforcement learning is more and more important in MAS (Multi-Agent System) research [7-9].

In this paper, role changing and reinforcement learning are combined to improve the learning ability of multi-robot coordination. Under centralize control framework, the distance nearest rule is presented, which means that the nearest robot ranges from obstacles is selected to be the master robot and control slave robots movement. Meanwhile, different from traditional way which reinforcement learning is directly applied in online learning of multi-robot coordination, behavior based multi-robot coordination is applied in this paper, the robot behavior weights are optimized through interacting with environment and the coordination policy based on maximum behavior value is used to plan robot collision avoidance behavior based on reinforcement learning behavior weight. The learning method developed in this paper is applied to the application related to collaboration portage of mobile robots. This is demonstrated by the simulation results presented in this paper.

2 Reinforcement Learning

Reinforcement learning RL [11], also called strengthen learning or encourage learning, is an important machine learning method. In dynamic environment, it can be used to acquire self-adaptability response behavior, which essence is an asynchronous dynamic planning method based on Markova Decision Process (MDP). Because of its ability of implementing the decision optimization through interacting with environment, RL has abroad application value in solving complex optimization control problem.

A general RL model is described in Fig.1. The agent and environment interact at each sequence of discrete time steps $t = 0,1,2,3,\ldots$. At each time step t, the agent receives some representation of the environment's *state*, $s_t \in S$, where S is the set of all possible states, and on that basis selects an *action*, $a_t \in A(s_t)$, where $A(s_t)$ is the set of actions available on state s_t. One step later, in part as a consequence of its actions, the agent receives a numerical *reward*, $\theta_{t+1} \in \Re$, and finds itself in a new state s_{t+1}.

At each time step, the agent implements a mapping from states to probabilities of selecting each possible action. This mapping is called the agent's *policy* and is denoted as π_t, where $\pi_t(s,a)$ is the probability that $a_t = a$ if $s_t = s$. RL methods specify how the agent changes its policy as a result of its experience. The agent's goal, roughly speaking, is to maximize the total amount of reward it receives in the long run.

$$Q(s,a) \leftarrow Q(s,a) + \alpha[\theta + \gamma \max_{a'} Q(s',a') - Q(s,a)]. \tag{1}$$

One of the most important RL methods is Q-Learning [12]. In this case, the learned action-value function, Q, directly approximates Q^*, the optimal action value function, independently of the policy being followed. In this setting, Q-values are updated according to Eq.1:

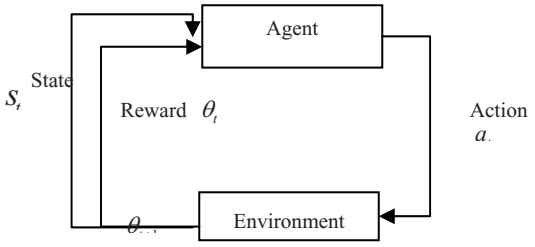


Fig. 1. The agent-environment interaction in reinforcement leaning

where, $Q(s,a)$ is the action value function of taking action a in state s, $Q(s',a')$ is the action value function in the resulting state (s') after taking action a, α is the step-size parameter between 0 and 1, θ is the reward, and γ is a discount rate. From Eq.1, it is clear that the agent only needs to consider each available action a in its current state s and choose the action maximizing $Q(s,a)$.

Table 1. Q-Learning: An off-policy Temporal Difference algorithm

Q-Learning
Initialize Q(s, *a)* arbitrarily Repeat (for each episode) Initialize s Repeat (for each step of episode) Choose a from s using policy derived from Q (e.g., $\varepsilon - greedy$) Take action a, observe θ , s' $Q(s,a) \leftarrow Q(s,a) + \alpha[\theta + \gamma \max_{a'} Q(s',a') - Q(s,a)]$ $s \leftarrow s'$ Until S is terminal

It may at first seem surprising that one can choose global optimal action sequences by reacting repeatedly to the local values of $Q(s,a)$ for the current state. This means the agent can choose the optimal action without ever conducting a lookahead search to explicitly consider what state results from the action. If action a is selected in state s , the value of $Q(s,a)$ for the current state and action summarizes in a single number is all the information needed to determine the discounted cumulative reward, that will be gained in the future.

Table 1 describes the Q-Learning algorithm. The agent's estimation of $Q(s,a)$ converges, limiting to the actual Q^* , provided the system can be modeled as a deterministic MDP, the reward function is bounded, and actions are chosen so that every state-action pair is often visited infinitely.

3 Behavior Weight Mechanism

3.1 Coordination Behavior of Multi-robot

In the multi-robot coordination work, we divide the process of coordination work into two steps：

1. Obtaining portage object, it contains three sub behaviors: (1) avoid obstacles (2) move to portage object (3) avoid portage zone
2. Portaging object: (1) avoid obstacles (2) move to target zone (3) enter portage zone

Shown as Eq.2 and Eq.3, the importance of every behavior is different in the coordination process：

$$\begin{aligned} AS_ACQUIRE_OBJECT=&Weight_{avoid}*MS_AVOID_OBSTACLES \\ &+Weight_{move}*MS_MOVE_TO_OBJECT \\ &+Weight_{home}*MS_AVOID_HOMEBASE , \end{aligned} \tag{2}$$

$$\begin{aligned} AS_DELIVER_OBJECT=&Weight_{avoid}*MS_AVOID_OBSTACLES \\ &+Weight_{bin}*MS_MOVE_TO_OBJECT_BIN \\ &+Weight_{home}*MS_AVOID_HOMEBASE . \end{aligned} \tag{3}$$

$Weight_{avoid,}$ $Weight_{move,}$ $Weight_{home}$ and $Weight_{bin}$ are the weights of avoiding obstacles, moving to portage objects, avoiding portage zone and moving to target zone respectively. MS_AVOID_OBSTACLES, MS_MOVE_TO_OBJECT, MS_AVOID_HOMEBASE and MS_MOVE_TO_OBJECT_BIN are the behaviors of avoiding obstacles, moving to portage objects, avoiding portage zone and moving to target zone respectively.

3.2 Behavior Weight Based on Reinforcement Learning

From section 3.1, it can get that robot coordination is composed of many behaviors which have different weights; also their importance are different. For automatic modifying the behavior on dynamic environment, we use reinforcement learning to automatic learn perfect weight combination. Robot can modify weight combination gradually and finish coordination portaging through interactive learning with environment.

In the learning initialization, every behavior weight has its own initialization value. During the learning process, we use evaluation function r to evaluate the weight according to the feedback information from environment. After the portaging process is finished, the weight will be recomputed, shown as Eq.4.

$$Weight_{new}=Weight_{old}+r . \tag{4}$$

If the portage is finished, r is positive number, the current weight is increased; else r is negative number, the current weight is decreased. The evaluation function is shown as Eq.8：

$$r=\begin{cases} \lambda & unfinished \\ \theta & finished \end{cases} , \tag{5}$$

where λ is negative number and θ is positive number.

4 Coordination Based on the Distance Nearest Rule

In the robot system, master-slave mechanism is implemented as the coordination infrastructure.

4.1 The Distance Nearest Rule

The definition of distance nearest rule is that selecting the robot which distance is nearest from obstacles in all coordination robots. The robot sends instruction message to control other robots action. Coordination portage is composed of obtaining object and portaging object, master-slave mechanism is only used in obtaining portage object. Shown as Eq.6 and Eq.7, in the process of portaging object, the robot which distance is nearest from obstacles is selected as the master robot which controls other robots movement.

$$\text{DistanceR}_{\min}=\frac{\min}{n\in N}Dis\tan ceR_n, \tag{6}$$

$$\text{R}_{\text{main}}=\text{R}_n \text{ if DistanceR}_{\min}=\text{DistanceR}_n. \tag{7}$$

where R_n is robot n, n=1,2,3,...,N, R_{main} is the master robot, $DistanceR_n$ is the distance from R_n to obstacle, $DistanceR_{min}$ is the nearest distance from robot to obstacles.

4.2 Coordination Policy Based on Maximal Behavior Value

In the robot coordination, its avoidance behavior is composed of acceleration and deflexion. We proposed coordination policy based on maximal behavior value.

The behavior value is shown as Eq.8

$$Value_tR = \gamma*(\alpha*Value_tS+\beta*Value_tD), \tag{8}$$

where $Value_tR$ is the behavior value of robot at time t, $Value_t$S is its acceleration value, α is the weight of acceleration value, $Value_t$D is the deflexion value of robot, β is the weight of deflexion value, γ is the weight of robot, the weight value of master robot is normal high than slave robots.

Shown as Eq.9, the coordination selecting policy is that master robot selects the avoidance behavior with the maximal behavior value from all robots and sends the behavior to other slave robots as their common avoidance behavior.

$$\pi_{t+1}=R \text{ if } Value_tR=\max_{R\in A}Value_tR, \tag{9}$$

where A is the set of all robots in the coordination work.

5 Simulation Experiment

For validating the efficiency of behavior weight method based on reinforcement learning, simulation experiment of multi-robot coordination portage system is carried out.

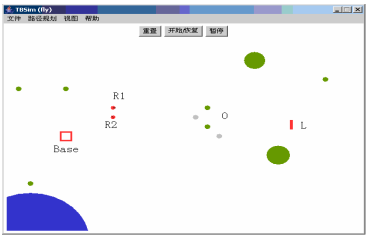


Fig. 2. Simulation experiment setting

Simulation environment is shown as Fig.2, where mobile robots R_1、R_2 and obstacles O are running in the area with 600*400. R's initiation movement velocity is 1.5; the initiation moving directions of two Rs are 0 and $3\pi/4$ denoted as polar coordinates, also dynamic changed with environment. O's velocity is 0. The object similar as long-pole is the portage object and its velocity is 0. The object similar as square zone is the portaging zone and its area is 10*10.

In the coordination work, the initiation weights of sub behaviors of obtaining object are set as $Weight_{avoid}$=1, $Weight_{move}$=1, $Weight_{home}$=1, $Weight_{bin}$ =1. The weight evaluation function of obtaining object is shown as Eq.10 :

$$r_{avoid}=\begin{cases}0 & \text{without collision}\\ 0.1 & \text{with collision}\end{cases},$$

$$r_{move}=\begin{cases}0 & \text{arrive at portage object}\\ 0.1 & \text{fail to get portage object}\end{cases}, \tag{10}$$

$$r_{home}=\begin{cases}0 & \text{avoid the portage zone}\\ 0.1 & \text{collide with portage zone}\end{cases}.$$

Meanwhile the initiation weights of sub behaviors of portaging object are set as $Weight_{avoid}$=1, $Weight_{move}$=1, $Weight_{home}$=1. The evaluation function of weight of portaging object is shown as Eq.11:

$$r_{avoid}=\begin{cases}0 & \text{without collision}\\ 0.05 & \text{collide with obstacles}\end{cases},$$

$$r_{bin}=\begin{cases}0 & \text{arrive at the destination zone}\\ 0.05 & \text{fail to arrive to the destination zone}\end{cases}, \tag{11}$$

$$r_{home}=\begin{cases}0 & \text{avoid the portaging zone}\\ 0.05 & \text{collide with portaging zone}\end{cases}.$$

Shown as Fig.3, on the condition of R、O、L and Homebase which is shown as Fig.2, collision can not be avoided when RL policy isn't be used. Shown as Fig.4-Fig.8, Rs select a successful path and finished the coordination work after three times RL learning through using RL policy.

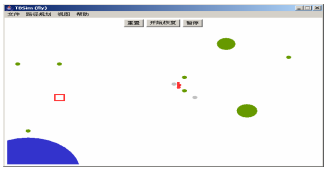


Fig. 3. Collide with obstacles

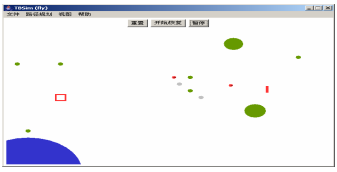


Fig. 4. Move to carrying object

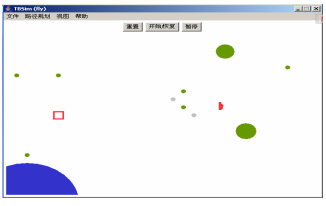


Fig. 5. Move to carrying zone

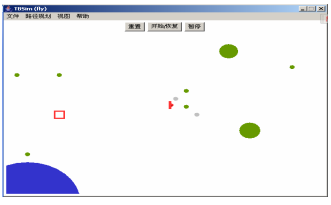


Fig. 6. Successful avoid obstacles

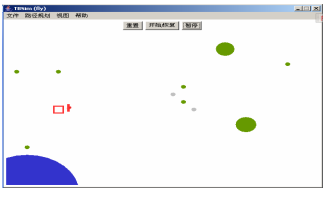


Fig. 7. Close to carrying zone

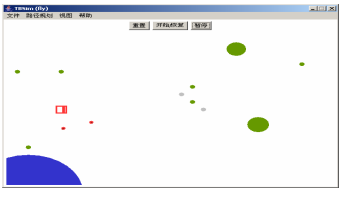


Fig. 8. Successful arrive at carrying zone

6 Conclusion

Multi-robot coordination is the hotspot problem in robot research area. Lots of academicians have do researches in multi-robot coordination area by using reinforcement learning. Normally, the main way of reinforcement learning in coordination is direct using it in on-line learning of robot. The paper combines role transformation and reinforcement learning. Under centralize control framework, the distance nearest rule is presented, which means that the nearest robot ranges from obstacles is the master robot to control slave robots movement. Meanwhile, different from traditional method which reinforcement learning is applied in online learning of multi-robot coordination, behavior based multi-robot coordination is applied in this paper, the robot behavior weights are optimized through interacting with environment and the coordination policy based on maximum behavior value is used to plan the collision avoidance behavior of robot based on reinforcement learning behavior weight. The learning method developed in this paper is applied to this application related to collaboration movement of mobile robots. This is demonstrated by the simulation results presented in this paper.

Acknowledgements

This work was supported by Shanghai Leading Academic Disciplines under grant T0103, the Excellent Young Teachers Program of Shanghai Municipal Commission of Education and Innovation Fund of Shanghai University.

References

1. Parker, L.E.: Lifelong Adaptation in Heterogeneous Multi-Robot Teams: Response to Continual Variation in Individual Robot Performance. Autonomous Robots 2000(8), 239–267 (2000)
2. Wang, Z., Hirata, Y., Kosuge, K.: An Algorithm for Testing Object Caging Condition by Multiple Mobile Robots. In: Proc. of the 2005 IEEEIRSJ International Conference on Intelligent Robots and Systems (IROS), Edmonton, Alberta, Canada, pp. 2664–2669 (2005)
3. Huntsberger, T., Pirjanian, P., Trebi-Ollennu, A.: Campout: A Control Architecture for Tightly Coupled Coordination of Multi-Robot Systems for Planetary Surface Exploration. IEEE Trans. on Systems, Man, and Cybernetics 33(5), 550–559 (2003)
4. Yamada, S., Saito, J.: Adaptive Action Selection without Explicit Communication for Multi-Robot Box-Pushing. IEEE Trans. on Systems, Man, and Cybernetics 31(3), 398–404 (2001)
5. Miyata, N., Ota, J., Arai, T., Asama, H.: Cooperative Transportation by Multiple Mobile Robots in Unknown Static Environment Associated with Real-Time Task Assignment. IEEE Trans. on Robotics and Automation 18, 769–780 (2002)
6. Minsky, M.L.: Theory of Neural Analog Reinforcement Systems and Its Application to the Brain Model Problem. Princeton University, New Jersey (1954)
7. Yang, E.F., Gu, D.B.: A Multi-Agent Fuzzy Policy Reinforcement Learning Algorithm with Application to Leader-Follower Robotic Systems. In: International Conference on Intelligent Robots and Systems, Beijing, China, pp. 3197–3202 (2006)
8. Guo, H., Liu, T.Y., Wang, Y.X., Chen, F., Fan, J.M.: Research on Actor-Critic Reinforcement Learning in RoboCup. In: Proceedings of the 6th World Congress on Intelligent Control and Automation, Dalian, China, pp. 9212–9216 (2006)
9. Duan, Y., Liu, X.G., Xu, X.H.: Design of Robot Fuzzy Logic Controller Based on Reinforcement Learning. Journal of System Simulation 18(6), 1597–1600 (2006)
10. Alur, R., Coucoubetis, C., Henzinger, T.A., HO, P.H., Nicollin, X., Olivero, A., Sifakis, J., Yovine, S.: The Algorithmic Analysis of Hybrid Systems. Theoretical Computer Science 138, 3–34 (1995)
11. Richard, S.S., Andrew, G.B.: Reinforcement Learning: An Introduction [M]. The MIT Press, Cambridge (1998)
12. Watkins, C.J.: Learning from Delayed Rewards. Ph.D. thesis, Cambridge University (1989)

An Algorithm of Coverage Control for Wireless Sensor Networks in 3D Underwater Surveillance Systems

Feng Chen, Peng Jiang*, and Anke Xue

Institute of Information and Control, Hangzhou Dianzi University, Zhejiang, China
pjiang@hdu.edu.cn

Abstract. In this paper, an algorithm of coverage control is introduced for three-dimensional underwater wireless sensor networks based on uniform deployment on surface which is suitable for water environment monitoring of wetland. Sensor nodes are uniformly deployed on surface of monitoring waters with equal interval. After deployment, sensors are lowered to appropriate depths selected by the distributed algorithm such that the maximum three-dimensional coverage of underwater sensor space is maintained. For the maximization of comprehensive index composed of coverage and average distance, simulation results reveal that our scheme outperforms two- and one-dimensional random deployment and depth random deployment under different network scale.

Keywords: Underwater sensor networks, 3D coverage, Coverage control.

1 Introduction

With the aggravation of global water pollution, there has been a growing interest in monitoring the marine environment for scientific exploration, commercial exploitation and coastline protection. Marine environment monitoring is restricted by unmanned exploration, localized knowledge acquisition and large scale of waters [1]. Wireless Sensor Networks (WSNs) is a new information acquisition technology with the development of wireless communication, embedded computing technology, sensor and MEMS technology (see[2] and [3]). Distributed and scalable Underwater Wireless Sensor Networks (UWSNs) provides effective settlement for water environment monitoring significantly different from any ground-based sensor network: Acoustic modem should be used for wireless communication equipment in underwater environment, and majority of underwater sensor nodes are mobile due to water currents, dispersion and shear processes. Furthermore, UWSNs is also different from existing small-scale Underwater Acoustic Network (UAN(see [4]and [5])): UWSNs relies on localized sensing and coordinated networking amongst large number (potentially hundreds to thousands) of low-cost and densely-deployed sensors. The extensive application prospects and high technical requirements make UWSNs become a new challenge.

To guarantee the integrity and accuracy of the overall information, the WSNs system first needs to determine whether the interest region is covered by a group of given sensor nodes or not. For example, the coverage of sensor network to some heavy disaster areas

* Corresponding Author.

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1206–1213, 2008.
© Springer-Verlag Berlin Heidelberg 2008

should be known in fire monitoring. Therefore, it is of great significance to have far and systemic research into the coverage problems of sensor network.Although many schemes at present have well solved the two-dimensional coverage control problem (see[6],[7],[8]and [9]), the current schemes of three-dimensional coverage control can only obtain approximate optimized results because it is a NP-problem [10]. How to design effective algorithms and protocols of coverage control for application requirement of WSNs in three-dimensional space is a meaningful subject to study. The coverage control for UWSNs is a typical three-dimensional coverage control problem. UWSNs is limited by floating mode mobility, limited acoustic link capacity and huge propagation delay. This new network model sets new request to coverage control[1]. [11] proposes an algorithm of coverage control for UWSNs based on random deployment on surface (sensors are randomly distributed in two dimensions of water surface, and it is named two-dimensional random deployment for short):Certain number of nodes are randomly placed on surface in the initial deployment. Each node then arranges its own depth by depths of neighbor nodes in its coordination space such that sufficient coverage is maintained. It accords with the characteristic of stronger fluidity of larger-area waters in ocean, river and lake due to the randomness of the distribution of nodes on surface, but cannot change the horizontal position of node after initial deployment. Besides, the random distribution of nodes inevitably causes that nodes are concentrated in certain regions, but seldom in other regions, which limits the three-dimensional coverage of underwater space for the entire sensor network.

Wetland waters has its own particularity compared to vast size of water area such as ocean, river and lake. The entire wetland water environment is divided into a large number of relatively independent small waters which are distributed widely with irregular shape and different size[12]. Based on [11], we introduce an algorithm of coverage control for UWSNs based on uniform deployment on surface (uniform deployment for short): Sensor nodes are uniformly placed on surface in the initial deployment. Their relative locations to each other normally do not change after deployment. We can merely arrange the depths of nodes to maximize the comprehensive index composed of coverage and average distance. Current in the partial small waters of wetland is gentle, so the motion of node on surface due to water currents, dispersion and shear processes can be ignored. Therefore, the uniform deployment can be applied to wetland water environment

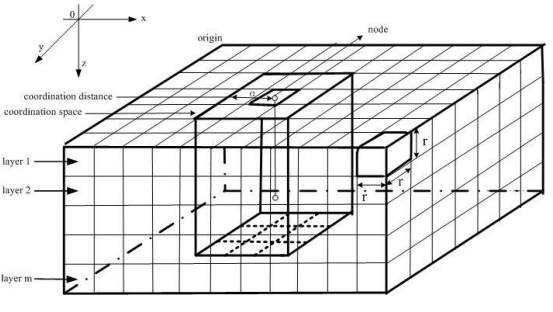


Fig. 1. The coordinate system for three-dimensional underwater wireless sensor networks (UWSNs)

real-time monitoring system based on WSNs. For the optimization objective of coverage control method that maximization comprehensive index, uniform deployment outperforms two-dimensional random deployment under different network scale.

The remainder of the paper is organized as follows: In Section 2, we discuss the theory and realization of the algorithm of uniform deployment; Through simulation, the comprehensive index, coverage and average distance of two- and one-dimensional random deployment, as well as uniform deployment are compared and analyzed in Section 3; And the paper is concluded in Section 4.

2 Algorithm of Uniform Deployment

2.1 Three-Dimensional Underwater Sensor Space

We model the sensor space as a cuboid and its coordinate system as shown in Fig.1. We accept an appropriate corner of the sensor space as the origin, X axis and Y axis as two dimensions in horizontal plane, Z axis as the direction vertically downwards the horizontal plane. We partition the whole sensor space into cubes *r* in size. *r* is called the resolution distance which is the edge length of a unit cube that sensor space is partitioned into. We use another parameter α which is called coordination distance. The space formed by a square along Z axis called coordination space, and the square is focusing on the node with its edge length of 2 α .The coordinates system of coordination space should satisfy the following conditions:

$$\begin{cases} x_i - \alpha \leq x_n \leq x_i + \alpha \\ y_i - \alpha \leq y_n \leq y_i + \alpha \end{cases} \tag{1}$$

Where x_n and y_n are coordinates of a given cube and x_i and y_i are the coordinates of the node. While a node arranges its depth, it exchanges information only with the neighbor nodes within its coordination space. Along with Z axis, all the cubes at the same depth construct a layer, e.g., all the cubes between depth 0 and depth *r* construct layer 1. There are m layers totally.

2.2 Coverage Model and Optimization Objective

We suppose the current in partial small waters of wetland is gentle and sensors are deployed in surface buoys whose relative locations to each other do not change after deployment. So we merely need to place the nodes to appropriate depths. Besides, Our scheme needs to be distributed, not relying on a central node that knows the up to date locations of nodes and arranges the depth of each node centrally.

Performance evaluation standard of coverage control for WSNs is very important to the analysis of the usability and validity of the coverage control strategy, and of which the coverage of monitoring region for network is of the most importance(see [13]). The coverage control hereafter for UWSNs has the following two basic goals in [11]:

1) Maximize coverage δ

$$\delta = \frac{n_c}{n_n}. \tag{2}$$

Where n_c is the number of cubes covered by at least one sensor and n_n is the total number of nodes in the underwater sensor space. Coverage here means that the cube is entirely within a sphere, of which the radius equal to the sensing range of node.

2) Maximize average distance θ

$$\theta = \frac{\sum_{i=1}^{k} e_i}{k}. \tag{3}$$

Where e_i is the distance between two nodes in node pair *i*. If there are n_n nodes in a sensor field, we can have as many as $k = \frac{n_n \cdot (n_n - 1)}{2}$ node pairs. Maximizing average distance aims to reduce the probability that all nodes are concentrated on certain depths.

Taking both coverage and average distance into consideration, we introduce the optimization objective of coverage control as follows:

$$\max S = a\delta + b\theta. \tag{4}$$

Where *S* is called comprehensive index, *a* and *b* are weight coefficients of coverage δ and average distance θ respectively.

2.3 Description of Algorithm

We introduce a heuristic algorithm to solve the coverage control problem for UWSNs. Each node exchanges information with neighbor nodes in its coordination space to arrange its depth. Uniform deployment takes the following steps:

(1) Model the sensor space as a cuboid and establish its three dimensional coordinate system. Uniformly deploy a certain number of nodes on water surface (Nodes are deployed with equal interval along X and Y axis), determine the coordinates along X, Y and Z axis. Initialize coordinate axis Z of each sensor node as 0;
(2) Determine each node's coordination space;
(3) Adjust the depth of each node by the number and depths of its neighbor nodes;
(4) Calculate the comprehensive index using the depths of nodes, determine the depths when the comprehensive index is maximized is the finial depths.

The adjustment of depths of each node is as follows:
Partition the depth of the water region into m layers. For each node,

1) If it doesn't have any neighbor nodes in its coordination space, then it is deployed in layer 1;
2) if it has only one neighbor node whose depth is in the upper half of the entire monitoring space, then the node is deployed in layer m;
3) if it has only one neighbor node whose depth is in the lower half of the entire monitoring space, then the node is deployed in layer 1;
4) if it has more than one neighbor nodes, then the node is deployed in the middle layer of the maximal and minimum depth.

In order to compare with two-dimensional random deployment and uniform deployment, we also introduce an algorithm of coverage control for UWSNs based on one dimensional random deployment and the other dimensional uniform deployment (Nodes are randomly distributed in one dimension and uniformly distributed in the other dimension of space. This scheme hereafter is called one-dimensional random deployment): Without loss of generality, nodes are randomly distributed in X axis and uniformly distributed in Y axis with equal interval in the initial deployment. Then arrange their depths with the same method as uniform deployment. Comprehensive index, coverage and average distance of two-dimensional and one-dimensional random deployment, as well as uniform deployment are analyzed and compared in the simulation.

3 Simulation Examples

Various numbers of sensor nodes are deployed over a space of $60 \times 50 \times 30$ in size. It is assumed that the sensing radius of each sensor node is 14, both the resolution distance r and coordination distance α are 10. Consider two limiting cases of comprehensive

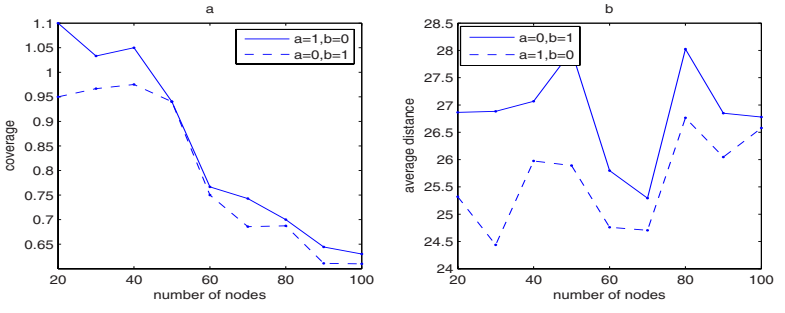


Fig. 2. Coverage and average distance of one-dimensional random deployment and uniform deployment

index: (1) take a=1, b=0, namely maximizing coverage; (2) take a=0, b=1, namely maximizing average distance. For one-dimensional random deployment, the coverage in case (1) is larger than in case (2) in Fig.2.a, while the average distance in case (2) is larger than in case (1) in Fig.2.b; The same conclusion can be obtain for the coverage and average distance of uniform deployment. It can be seen that if coverage of the comprehensive index is maximized, then the average distance can't be maximized. Also the maximization of average distance can't ensure the maximization of coverage.δ is much smaller than θ , so the weight coefficients should be appropriately chosen in the simulation. We take a=100 and b=1.

In Fig.3 the comprehensive index of two- and one-dimensional random deployment and uniform deployment are shown for varying numbers of nodes deployed over the monitoring space. When the number of nodes increase, the comprehensive index of the three schemes all decrease. With the same number of nodes, the comprehensive index of the uniform deployment is higher than two- and one-dimensional random deployment. Because with certain quantity of nodes, the more uniformly the nodes are distributed,

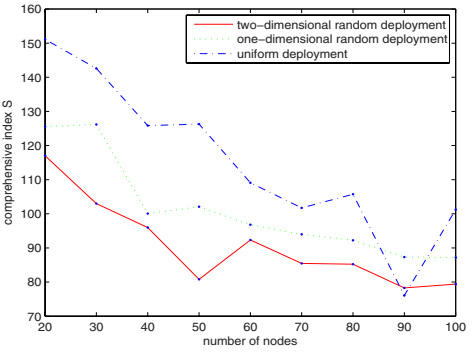


Fig. 3. Comprehensive index S for varying numbers of nodes when a=100, b=1

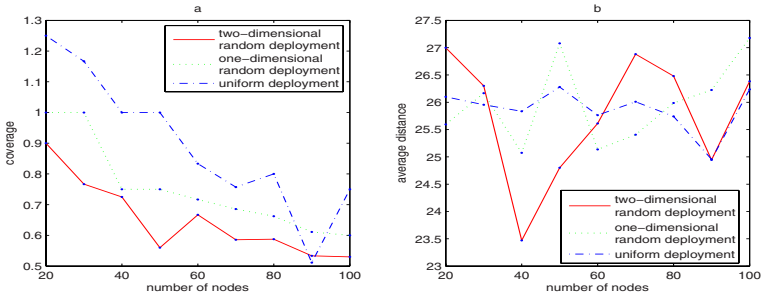


Fig. 4. Coverage δ and Average distance θ for varying numbers of nodes when a=100, b=1

the larger the coverage of the monitoring region is, thus the larger the comprehensive index is. Uniform deployment can perform the most uniform distribution of the three methods, while the situation that nodes are concentrated on certain regions and rare on other regions inevitably exists in two- and one-dimensional random deployment. The comprehensive index S of uniform deployment suddenly drops when the number of nodes is 90, which is even smaller than two- and one-dimensional random deployment. It is related to the distributed location of nodes on surface. In the simulation, 90 nodes are uniformly distributed on surface by 9×10 pattern with 9 nodes distributed on each column. Some nodes exactly are in the edge of the cube, so these nodes cover less cubes, which causes the drop of coverage and comprehensive index. The fluctuation of the comprehensive index curves of the two- and one-dimensional random deployment is caused by the randomness of the deployment of the nodes on surface.

We analyze the coverage and average distance withdrawn from the comprehensive index in Fig.3 respectively. It can be seen from Fig.4.a that coverage δ of two- and one-dimensional random deployment and uniform deployment decrease with the increase of number of nodes deployed in monitoring region. Because the coverage is defined as the rate of the number of cubes covered by at least one sensor to the total number of nodes in the space, and denominator increase more quickly than numerator with the increase

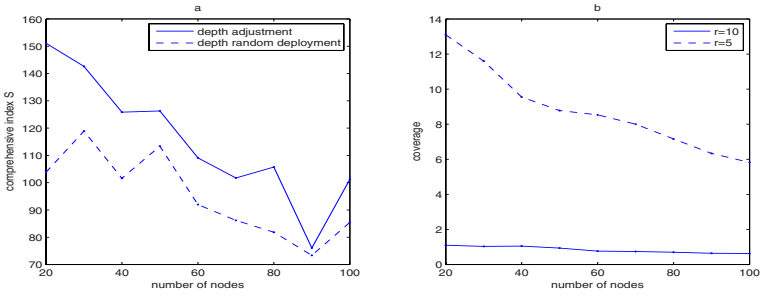


Fig. 5. a. Comprehensive index S for varying numbers of nodes. b.Coverage δ of one-dimensional random deployment under various resolution distance r for varying numbers of nodes, and a=1, b=0.

of number of nodes.With the same number of nodes, the coverage δ of the uniform deployment is larger than two-dimensional and one-dimensional random deployment with the same reason above. The change tendency of coverage with the varying numbers of nodes is consistent with the change tendency of comprehensive index S. Because the weight coefficient a of coverage is large in the comprehensive index S, so the change tendency of coverage has basically decided the change tendency of comprehensive index. From Fig.4.b, We can see that there is no obvious closely relationship among the values of average distance of the three methods. Because the value of weight coefficient b of average distance is small, and in the process of maximization of comprehensive index, it is unable to maximize the average distance. The randomness of distribution of nodes in two- and one-dimensional random deployment causes the large fluctuation of average distance curve.

The comprehensive index of two different deployment schemes is compared in Fig.5.a. One method is uniformly distributing nodes on water surface with equal interval in the initial configuration with random depth selection, which is called depth random deployment for short. The other one is the uniform deployment called depth adjustment as comparison. With the same number of nodes, the comprehensive index of depth adjustment is obviously larger than depth random deployment.

Coverage of one-dimensional random deployment for varying resolution distance r is shown in Fig.5.b for various numbers of nodes. r increases, and the volume of unit cube also increases. Therefore, when the radius of sensor node keep invariable, the larger r is, the less the cubes covered by at least a sensor are, and the smaller the corresponding coverage is. We can see from Fig.5.b that with the same number of nodes deployed, the coverage when r=5 is far larger than r=10. r should be chosen prudently according to the sensing radius of sensor node in the application to construct three-dimensional submarine sensor network.

4 Conclusions

We propose an algorithm of coverage control for UWSNs based on uniform deployment on surface (uniform deployment) for the relative stability of wetland water environment.

This algorithm may be operated on each sensor node. Deploy nodes uniformly on water surface at first, and then adaptively arrange the depth according to the locations of its neighbor nodes in the coordination space to maximize the comprehensive index composed of coverage and average distance such that the maximum coverage of the underwater three-dimensional space can be obtain. Two- and one-dimensional random deployment, depth random deployment and uniform deployment are compared. The simulation results indicate the comprehensive index of uniform deployment surpasses the first three methods under different network scale.

Acknowledgment. This paper is supported by the National Natural Science Foundation of China (NSFC-60604024), the Important Project of Ministry of Education (207044), the Key Science and Technology Plan Program of Science and Technology Department of ZheJiang Province (2008C23097) , the Scientific Research Plan Program of Education Department of Zhejiang Province (20060246) and the Sustentation Plan Program of Youth Teacher in University of Zhejiang Province (ZX060221) .

References

1. Ren, F.Y., Huang, H.N., Lin, C.: Wireless Sensor Networks. J. Software Journal 14, 1282–1291 (2003)
2. Cui, J.H., Kong, J.J., Marion, G., Zhou, S.L.: Challenges: Building Scalable and Distributed Underwater Wireless Sensor (UWSNs) for Aquatic Applications. Technical Report, UCONN CSE (2005)
3. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y.: Wireless Sensor Networks: A Survey. J. Computer Networks 4, 393–422 (2002)
4. Sozer, E.M., Stojanovic, M., Proakis, J.G.: Undersea Acoustic Networks. J. IEEE Journal of Oceanic Engineering 25, 72–83 (2000)
5. Proakis, J.G., Sozer, E.M., Rice, J.A., Stojanovic, M.: Shallow Water Acoustic Networks. J. IEEE Communications Magazine 39, 114–119 (2001)
6. Megerian, S., Koushanfar, F., Potkonjak, M., Srivastava, M.B.: Worst and Best-case Coverage in Sensor Networks. J. IEEE Trans. on Mobile Computing 4, 84–92 (2005)
7. Meguerdichian, S., Koushanfar, F., Qu, G., Potkonjak, M.: Exposure in Wireless Ad-hoc Sensor Networks. In: ACM Int'l Conf. on Mobile Computing and Networking (MobiCom)., pp. 139–150. ACM Press, New York (2001)
8. Meguerdichian, S., Koushanfar, F., Potkonjak, M., Srivastava, M.B.: Coverage Problems in Wireless Ad-hoc Sensor Network. In: IEEE INFOCOM, pp. 1380–1387. IEEE Press, Anchorage (2001)
9. Li, X.Y., Wan, P.J., Ophir, F.: Coverage in Wireless Ad hoc Sensor Networks. J. IEEE Trans. on Computers 52, 753–763 (2003)
10. Huang, C.F., Tseng, Y.C.: A Survey of Solutions to the Coverage Problems in Wireless Sensor Networks. J. Journal of Internet Technology 6, 1–8 (2005)
11. Erdal, C., Hakan, T., Yasar, D., Vadat, C.: Wireless Sensor Networks for Underwater Survelliance Systems. J. Ad Hoc Networks 4, 431–446 (2006)
12. Jiang, P.: Survey on Key Technology of WSN-based Wetland Water Quality Remote Realtime Monitoring System. J. Chinese Journal of Sensors and Actuators 20, 183–186 (2007)
13. Ren, Y., Zhang, S.D., Zhang, H.K.: Theories and Algorithms of Coverage Control for Wireless SensorNetworks. J. Journal of Software 17, 422–433 (2006)

An Embedded SoC System IP to Trace Object and Distance

Eun-Tack Oh, Kwang-Sung Jung, Sung-Min Lee, and Cheol-Hong Moon

52 Hyodeok-Ro, Gwangju University, Korea
silvertack5@nate.com,
magi8051@nate.com,
cruzdxx@naver.com,
chmoon@gwangju.ac.kr

Abstract. This study created an An Embedded SoC System IP to Trace Object and Distance. The moving object tracing system uses video input, image separation of moving objects and block-matching. The distance detection system receives stereo video input through 2 CCD cameras. Using a decoder, the image is changed to the YCbCr4:2:2 for- mat and only the Y signal is saved in the 4*256*8bit shift register of the Dual-Port SRAM. When the fourth line of the axis (256 pixels per line) data of each video is saved in the shift register, each standard block and block-matching are set and stereo-matching is carried out. The matching range is set to the 100 pixels to the right of the standard block coordi- nates. As a result of the matching procedure, the Depth value, which is the distance information, is saved in SRAM, and the Depth Map is made and output to the TFT-LCD screen.

abstract SoC system.

Keywords: Embedded SoC, IP Develop, Image Process, CCD Cameras.

1 Introduction

With recent rising interest in security and crime prevention the existing personal computer-based system is being set aside to make room for adaptation of an embedded system. However, the existing real time handling system complex, so design is difficult and the system is big. If such image processing system is made using hardware logic, real time processing can be possible. This study designed a simple moving object tracing system using software, and used hardware language VHDL to design a logic IP for the distance detection system, to build an embedded SoC system.

2 Moving Object Tracing System

Figure 1 shows the block diagram for the moving object tracing system. The signal entered through the CCD camera using the frame grabber board, and

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1214–1221, 2008.
© Springer-Verlag Berlin Heidelberg 2008

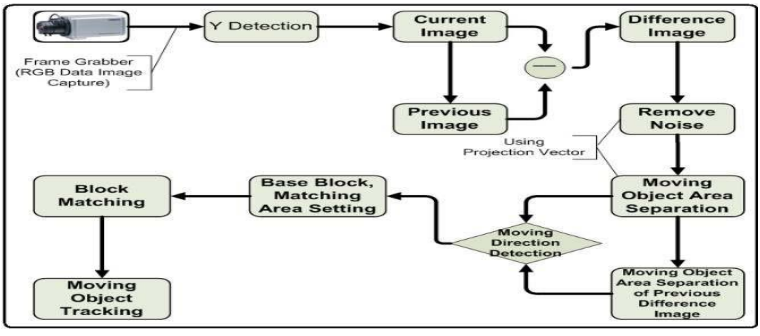


Fig. 1. Moving object tracing system

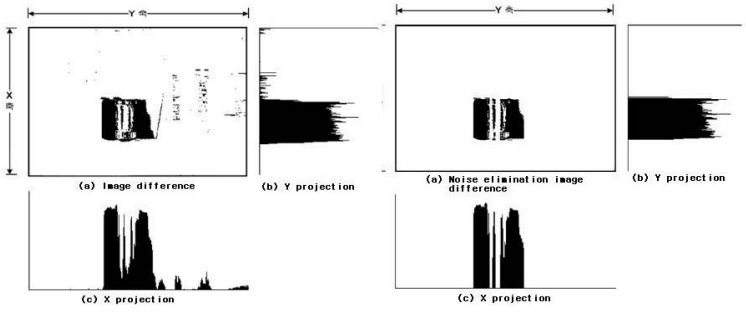


Fig. 2. The projection of image difference and Elimination of noise

transform to RGB only extracts and saves the Y signal. input for one frame is the pixel differences between the current image and the previous image is used to find the image difference. Small noises of the image difference are eliminated using a projection vector, and moving objects are separated. The currently separated moving object area and the previous area are compared to find the direction of movement. Standard block and matching area are set according to the direction of movement, and block matching is carried out to trace the moving object. Figure 2 shows the noise elimination results using the projection vector in Figure right. For noise elimination, only values that are under 1/3 of the highest value of the projection vectors of each the X and Y axis were eliminated.

3 Distance Detection System

Figure 3 shows the block plan of the distance detection system. The analogue signals of 2 CCD cameras are converted to YCbCr4:2:2 format through an image decoder, and only the Y signals are extracted. Next, the Dual-Port SRAM Control block is used to save the image in each Dual-Port SRAM. When input

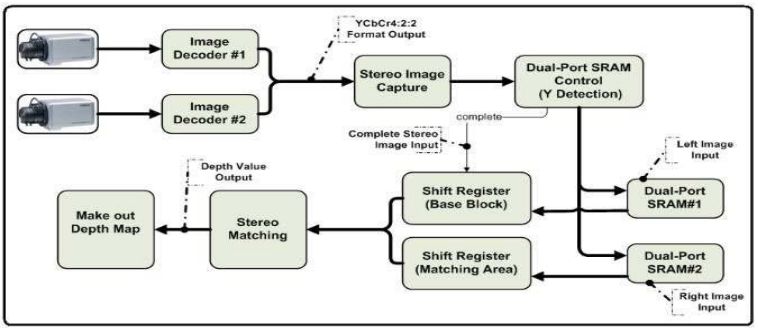


Fig. 3. Distance detection system block diagram

for one video is complete, a inputted signal is sent through the Dual-Port SRAM Control block. The shift register saves each of the Dual-Port SRAM data in the shift register simultaneously with the acknowledge signal. When the forth line (256 pixels per line) data of each video is saved in the shift register, each standard block and block-matching are set and stereo-matching is carried out. As a result of matching, the Depth value, which is the distance information, is saved in SRAM, and the Depth Map is made and an output is made on the TFT-LCD screen.

4 Distance Detection System IP Block

Figure 4 shows the distance detection system IP block. The system is largely divided into processor are, PLD area and external hardware area. The processor area consists of ARM922T 32bit RISC core and Strip-to-PLD Bridge to interface the PLD area, and the PLD-to-Stripe Bridge to interface the processor area from the PLD area, and the internal Dual-Port SRAM. The PLD area consists of the

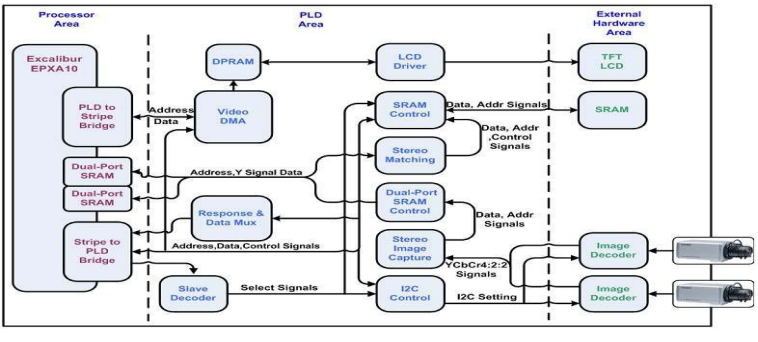


Fig. 4. Distance detection system IP block

I2C Control, stereo video capture Dual-Port SRAM control, stereo-matching, SRAM Control, DMA, LCD Driver IP,Booting is done in order to set the internal register of the processor's image decoder, 157 words are sent to the I2C Control block through the Stripe-to-PLD Bridge, and the transmitted data set the image decoder's internal register. When setting is complete, the image decoder outputs YCbCr4:2:2 format data. After completing settings, stereo image capture blocks are used to simultaneously receive 2 image decoder data and the Dual-Port SRAM Control block is used to save in the Dual-Port SRAM 1 and 2, each. When input for one frame is complete, the acknowledge signal is sent to the stereo-matching block. The stereo-matching block consists of two 4*256*8bit shift register, and each saves one image. When all 4 lines of the X axis of the image are saved, the 4*4 standard block and matching block are set to carry out stereo-matching, and the result are saved in SRAM using the SRAM Control block. When stereo matching for one frame is complete, the Depth Map is finished. Once finished, the processor area's IRQ is made to map SRAM and SDRAM and the DMA block and LCD Driver block are passed through to output the Depth Map on the TFT-LCD screen.

5 Test and Results

Figure 5 shows the moving object tracing system flow chart. When the program starts, RGB image is received through the frame grabber and the Y signal is extracted. The current Y image and previous video's pixel difference is used to find the image difference and check if there is a moving object, and if there is one, noises of the image difference are eliminated using a projection vector, and the moving object is separated. When separation is complete, the moving object separation coordinates of the former image are compared to check the direction of movement, and a standard block of the moving object in the current image is set. In order for precise block-matching the current image and previous image edges are extracted and block-matched to trace the moving object. ((a): Start stereo

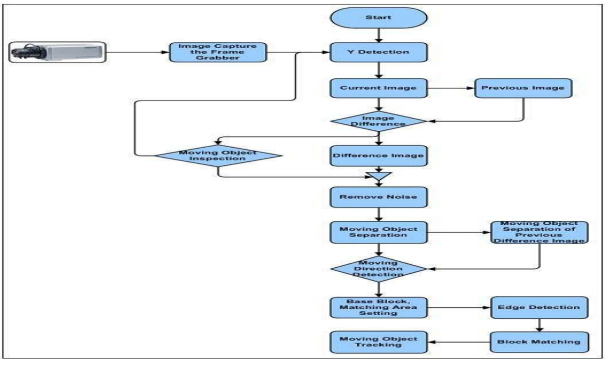


Fig. 5. Flow chart of moving object tracing system

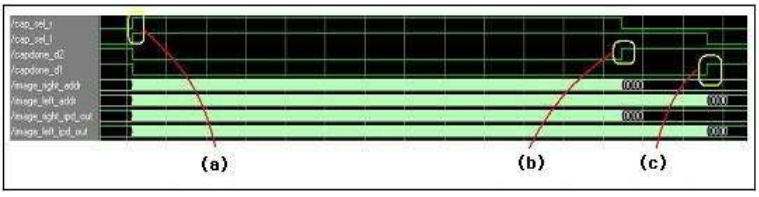


Fig. 6. Stereo image capture simulation

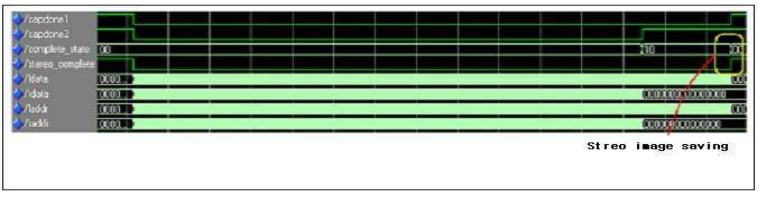


Fig. 7. Dual-Port SRAM control simulation

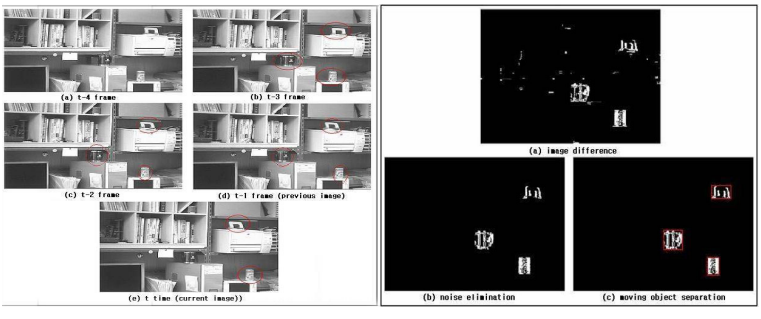


Fig. 8. Time difference image and Noise elimination and moving object separation

image capture, Right (b): Completed capturing right image, Lef (c): Completed capturing left video) Figure 6 is the simulation results of the stereo image capture IP. The start of the stereo image capture uses the i2c-set (image decoder internal register setting completion signal) and stream code extraction signals to confirm the cap-sel-r, cap-sel-l, which is the video capture activation signal. When image of a frame from each image capture blocks are received, the capdone-d1, capdone-d2 signal is output as '1'. Figure 7 shows the simulation results of the Dual- Port SRAM Control IP. While the signal of capdone1, capdone2 in the stereo image capture IP is '0', each image data is saved in the Dual-Port SRAM. When both signals are '1', the complete signal is changed to '1' and the stereo matching block is sent a completed signal. Figure 8 shows the image difference for tracing moving objects(left). Pictures (b), (c), (d) show 3 objects moving, and in (e), 2 objects have moved. This study used the continuous video in chapter 4 to trace

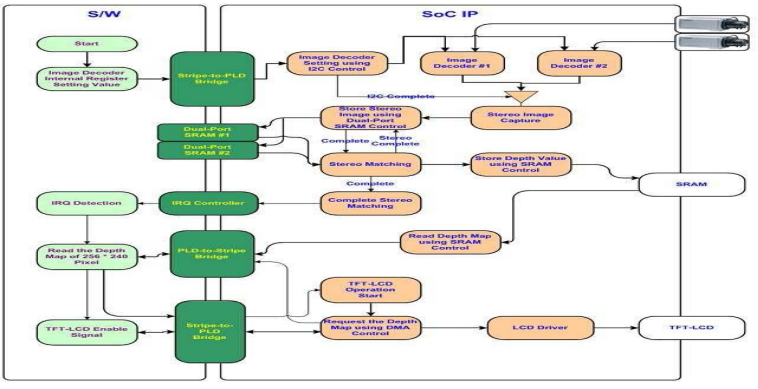


Fig. 9. Distance detection system flow chart

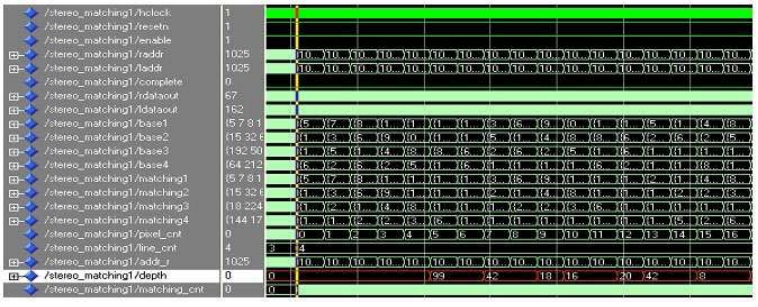


Fig. 10. Stereo-matching simulation

moving objects. elimination of noise through a projection vector on the image difference, and separation of the moving object(figure 8 right). (a) is the image difference of the t-4 frame and the t-3 frame, and includes 3 moving objects and noise. (b) has eliminated the noise from the image difference in (a) using a projection vector, and (c) shows three divided areas of moving objects through the projection vector. Figure 9 shows the distance detection system flow chart.

The distance detection system is divided into the ARM922T based software part and the SoC IP part. The software part goes through the Strip-to-PLD Bridge to set the image decoder and internal register value, and I2C Control block, and the internal register of the image decoder is set. When setting is complete, the YCBCr4:2:2 format data is output through the image decoder, stream code is inspected and image input is obtained. The stereo image capture part receives signals from the two image decoders and uses the Dual-Port SRAM Control to save each when one frame is saved, the acknowledge signal is sent to the stereo-matching block. Dual-Port SRAM data is saved in the shift register in the stereo-matching block, and the standard block and matching block are set and stereo-matching is carried out. When stereo-matching is completed in the

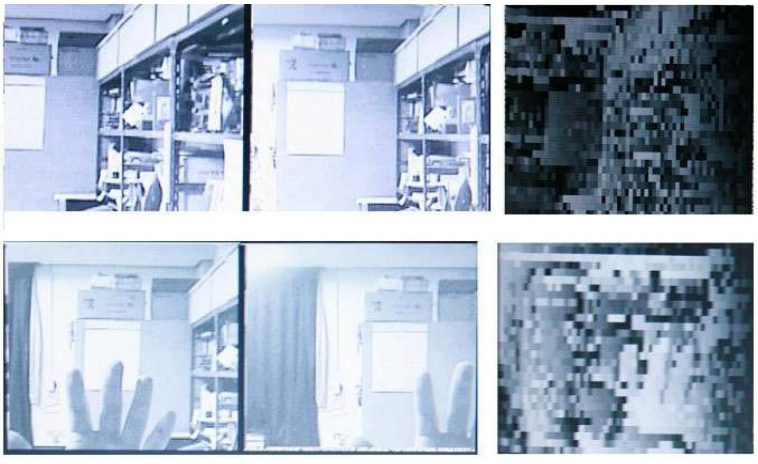


Fig. 11. Depth map output results

matching area, the result Depth value is saved in SRAM using SRAM Control. When stereomatching for one frame is complete, an IRQ signal is sent and Depth Map data of the SRAM, SDRAM is read and the output appears on the TFT-LCD screen using the DMA Control block and LCD Driver block.

Figure 10 shows the results of stereo-matching. Input data was entered to view simulation results. The same data was entered in the beginning part, and later, different data was entered. Data was only saved in the shift register until before the line-cnt (X axis line) was at 4 (the start of the fifth line of the X axis), and after 4, stereo-matching was carried out. Stereo-matching output results showed the standard block starting address, which is the SRAM address, and the Depth value of the stereo-matching results.

Figure 11 shows the Depth Map output results obtained from the stereo image using the I2C control, stereo capture, Dual-Port SRAM Control, stereo-matching, SRAM Control and TFT-LCD IP. also shows the output results when a hand is placed in front of the cameras. This study found the CCD camera's focal distance according to the test values.

6 Conclusion

The distance detection system used a stereo image and saved each in the Dual-Port SRAM, and when one frame was saved, the Dual-Port SRAM data needed for stereo matching was saved in the shift register. The shift register's standard block where the Dual-Port SRAM 1 data was stored, and the 2 matching block were set and stereo-matching was carried out. The distance detection system was implemented as an IP on the SoC Chip of the embedded SoC board. The implemented IP was an I2C Control to set image decoder, Stereo Image Capture to receive stereo images, Dual-Port SRAM Control to save stereo videos in

Dual-Port SRAM, Stereo-matching, SRAM Control to save SRAM result values, and TFTLCD IP to have a TFT-LCD screen output. This study expanded the existing single moving object tracing system to a multi-moving objects tracing system, and achieved clear results. The implemented distance detection system using stereo-matching operated at 7.2ms, and real time processing was made possible. The stereo matching results of this study show that there are many errors take place because of occlusion area and impulse noise. Also, impulse noise occurs in input videos. If algorithms and circuits are added in the design to solve these two errors, a more precise distance detection system could be implemented. Also, moving object tracing system and distance detection system were separately designed, but if the two systems are collaborated, it could create a more effective system.

Acknowledgments. This research was financially supported by the Ministry of Commerce, Industry and Energy(MOCIE) and Korea Industrial Technology Foundation(KOTEF) through the Human Resource Training Project for Regional Innovation, and Korea Institute of Industrial Technology Evaluation and Planning(ITEP) through the Technical Development of Regional Industry.

References

1. Park, H.-s.: Real Time Moving Objects Tracing and Distance Measurement System using Active Stereo Camerass (2006)
2. Kim, H.-O.: Real Time Moving Object Tracing System Based on SoC System (2006)
3. Bae, J.-b.: Embedded System for Extracting Subjects using Stereo Videos (2005)
4. Gyaourova, A., Kamath, C., Cheung, S.-C.: Block Matching for Object Tracking (2003)
5. Han, J.-h., Lee, E.-j.: Practical Observation Equipment Execution through Real Time Video Analysis. Hanbat University Specialist Graduate School of Information Communication (2002)
6. Kang, D.-j., et al.: Digital Video Treatment using Visual C++. Cytech Media (2003)

RSS Based Localization Scheme Using Angle-Referred Calibration in Wireless Sensor Networks

Cong Tran-Xuan[1], Eun-Chan Kim[2], and Insoo Koo[1]

[1] School of Electrical Engineering, University of Ulsan
[2] Department of Information and Communications,
Gwangju Institute of Science and Technology
txuancong@yahoo.com, tokec@gist.ac.kr, iskoo@ulsan.ac.kr

Abstract. In wireless sensor networks (WSNs), one of the conventional techniques for localization is the received signal strength (RSS) method where the RSS is utilized to get the distance measurement spatially between sensor nodes. However, in real localization system, the RSS is sensitively affected by many surrounding factors and tends to be unstable, so that it degrades accuracy in distance measurement. In this paper, we propose the angle-referred calibration based RSS method where angle relation between sensor nodes is used to perform the calibration for better performance in distance measurement. As results, the proposed scheme shows that it can provide a higher precision than the omni-direction calibration.

Keywords: Wireless sensor networks, angle-referred calibration, received signal strength, lateration.

1 Introduction

Wireless sensor networks (WSNs) have been attracting considerable research interest. WSNs have diverse applications in many fields such as military, environment, industry, agriculture and so on. Equipped with micro-integrated sensors, sensor nodes can sense a lot of real-time physical information such as humidity, temperature, velocity, etc. In most cases, the measurement data will be meaningless without knowing the location where the data are obtained. Moreover, the location of node can be used in queuing inquiry, inventory management, location-based routing, and so on. Subsequently, localization has a critical role in operation in WSNs.

There are four common methods of measuring distance [1] which are time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA) [2] and received signal strength (RSS) [3]. The TOA technique needs strict time synchronization of the whole sensor network and in realistic application it is very difficult to implement the time synchronization between beacon nodes and unknown nodes. This approach is therefore cost and difficult to deploy in real systems. Although the TDOA uses the difference of time arriving at sensor nodes to calculate the distance, it reduces the tight of time synchronization, but it also uses the time method that requires the synchronization. It subsequently takes cost and complexity to implement in real systems but the

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1222–1233, 2008.
© Springer-Verlag Berlin Heidelberg 2008

concession is that it obtains a good performance in distance measurement. There is therefore a clear trade-off between ranging error and complexity and cost. In AOA method, sensor nodes have to be equipped with an array to define the angle that the signal arrives from other sensor nodes. Thus, it is prominent in cost, size, and power consumption. On the other hand, the RSS is one of the simplest methods in distance measurement such that we can implement easily in real systems. This technique is based on a standard feature found in most wireless devices, a received signal strength indicator (RSSI). It is attractive because of no additional hardware, and is unlikely to significantly impact local power consumption, sensor size and cost. However, a difficulty to gain the performance [4] with this method is that, in the communication channel, the signal strength is impacted by some factors causing error. Thus, if we can well model the communication channel [6] and estimate the channel parameters, the RSS method is a prominent choice in many applications. To model the path loss of the signal, and further get the channel parameter suitable to a changing environment, the RSS self-calibration protocol was proposed in [5] where the RSS calibration is used to gain the distance measurement between nodes. However, this calibration is based on omni-direction and it does not reflect the affection from environment according to the direction. In working environment, the overall performance of RSS is dependable on the direction that one signal transmitted and it is not effective if we assign all with the same characteristics. In this paper, we therefore propose the scheme that uses the angle-referred RSS calibration. The calibration uses the direction of arrival basing on relation between distances in geometry and RSS between nodes. Finally, this angle-referred self-calibration is used to get the coarse distance measurement. By exploiting this aspect, the localization algorithm can minimize the error affects of mobility property of sensor nodes and dynamically changing environment.

The rest of paper is organized as follows. In section 2 we provide the basic propagation model in wireless communication channel. In Section 3, we propose the angle-referred calibration based RSS method. In section 4 we present the simulation results. Finally, we draw conclusions in Section 5.

2 Propagation Model

In free space model, the path loss of power is proportional inversely with the square of distance d between transmitter and receiver. Let us denote the received power by Pr(d). Then this received power Pr(d) is related to distance d through the Friis equation:

$$P_r(d) = P_t.G_t.G_r.\left(\frac{c}{4\pi fd}\right)^2, \tag{1}$$

where Pt is the transmitted power, Gt is the transmitter antenna gain, Gr is the receiver antenna gain, c is the propagation speed of radio frequency signal and f is the frequency of signal.

The free-space model however is an over-idealization because in the real environment, the propagation of signal is affected by reflection, diffraction, and scattering. Of course, these effects are environment dependent. Using the statistical model [6], this

model shows the relationship between the RSS and distance dependent on the environment systems work. The formula is expressed by,

$$P_r(d) = P_t \cdot \left(\frac{c}{4\pi f} \right)^2 \cdot \frac{1}{d^\eta}, \tag{2}$$

where η is the path loss exponent [5] according to environment, the gain of transmitting and receiving antenna is standardized to be 1. Table 1 shows the path-loss exponent value according to environments.

Table 1. Path-loss exponent in some environments

	Environment	η
Outdoor	Free Space	2
	Shadowed urban area	2.7 to 5
In building	Line-of-sight	1.6 to 1.8
	Obstruct	4 to 6

We can model the value of $P_r(d)$ at a distance d with reference to $P_0(d_0)$ at distance d_0 as the following formula:

$$\frac{P_r(d_0)}{d^\eta} = \frac{P(d) + X_\sigma}{d_0^\eta} \tag{3}$$

where Xσ is a normal random variable with zero mean and standard deviation σ that accounts for fading effects and noise in the environment.

3 Self- localization Algorithm

In the Section, we propose a RSS based localization scheme using angle-referred calibration where we use the angle-referred RSS calibration to estimate the distance between beacon nodes to unknown nodes, and further we apply the Least Square (LS) method to get the best estimation of location of unknown node. The proposed scheme can be divided into two processes:

- Distance measuring process: estimate the coarse distance between the nodes and the beacon nodes by angle-referred calibration.
- Location calculating process: calculate the location of unknown nodes by lateration method.

3.1 Distance Estimation Using Angle-Referred Calibration

In the paper, we assume that there are some beacon nodes that know their locations by equipped GPS or by artificially deployed at known position and lots of unknown nodes in the area covered by the range of these beacon nodes. Every node also has the capability of RSSI. The considered sensor networks are deployed on a plane, so we just need to locate the 2-D position of unknown node. We have N beacon nodes (xi,

yi) i from 1 to N, and M unknown nodes (xj, yj) j from 1 to M. We also assume that sensor nodes have their ID, and that each beacon node maintains a beacon message which contains its ID and location.

With these assumptions, the proposed distance estimation can be described as follows: When a beacon node receives beacon messages from its neighbor beacon nodes, it periodically computes the calibration parameter using the RSS and location of neighbor beacon nodes according to the angle formed with its neighbor beacon nodes. The details of this procedure will be explained in subsection 3.1.A and 3.1.B. The next procedure will be initiated by an unknown node. That is, when an unknown node needs to localize its position, it transmits a message that contains its ID to beacon nodes in its transmission range in order to make a request for the calibration parameter. When beacon nodes in range receive the message, they estimate the distance from this unknown node to them. After that, these beacon nodes transmit a message to neighbor beacon nodes that contains the distance and the ID of the unknown node. When a beacon node receives more than two messages from its neighboring beacon nodes, it takes only two sets of data from these beacon nodes that are not close, in order to calculate the direction of arrival from the unknown node. Then this beacon node transmits a packet containing the angle-calibration parameter in the sector that unknown node belongs to and the ID of that unknown node. Finally, the unknown node uses this parameter and the RSS to calculate the distance from it to the beacon node. The details of this procedure will be explained in subsection 3.1.C.

3.2 Defining the Angle of Arrival Relation between Beacon Nodes

In this subsection, the procedure to calculate the angle of arrival between beacon nodes is illustrated.

As showed in the Fig. 1, when a beacon node A receives beacon packets containing location of neighbor beacon nodes, in this case they are B and C. Beacon node A uses these data to calculate the angle of arrival from beacon node B and C. The angles between beacon nodes are oriented relatively to X-axis. The angle relation between beacon node A and B is determined by,

$$\hat{BA} = \tan^{-1} \frac{y_B - y_A}{x_B - x_A} \tag{4}$$

For a general expression, the (4) formula can be expressed by,

$$\beta = \tan^{-1} \frac{y}{x} ,$$

where β is defined as the angle between two beacon nodes, x is the difference of first component between two beacon nodes' coordinates and y is the difference of the second component.

Generally, the angle of arrival is in range (0:2π), and β is the angular function of x and y:

$$\beta(x, y) = \begin{cases} \tan^{-1}\left(\frac{y}{x}\right)(mod\, 2\pi), & x \geq 0 \\ \pi + \tan^{-1}\left(\frac{y}{x}\right), & x < 0 \end{cases} \tag{5}$$

Simultaneously, beacon node A also calculates the calibration parameter according to sector between node B and C. And in next subsection, the process to get the calibration will be explained.

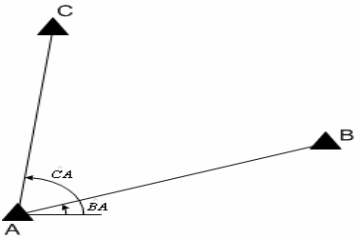


Fig. 1. Angle of arrival between beacon-nodes

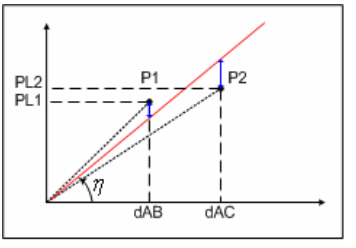


Fig. 2. An example of calculating calibration parameter

3.3 Calculating the Calibration Parameter

As we have known in the wireless communication channel, especially in the wireless sensor network, the received signal strength is affected by many factors mainly being fading and motion. The fading, the variation of the received power level as a function of time, compromises the accuracy of the estimation relying on power. Yet, the large number of measurement can resolve this error. Unfortunately, the sensor nodes, in some scenarios, can be movable, so it means that the accuracy of estimation will get worse when using a large number of samples. Moreover, the dynamic mobility characteristic makes the environment change overtime. So the calibration is used to adapt to the changing of environment surrounding sensor nodes.

From the formula of equation (2), we cast as followed,

$$\frac{P_t}{P_r} = \left(\frac{4\pi . f}{c}\right)^2 d^{\eta} \tag{6}$$

If we use the dB scale model, the relationship between the path loss of power and distance is seen linearly. Therefore, the (6) can be showed in dB scale,

$$P_{PL}[dB] = \eta . 10 \log d \tag{7}$$

where $P_{PL}[dB]$ equals $P_t[dB] - P_r[dB] - 20\log\left(\frac{4\pi . f}{c}\right)$, the path loss of power is proportional to $10\log d$, so if we can find the best line of the relationship, it is a calibration of RSS. Our process is finding the best parameter η that is the slope of the line closely describing the relationship. In the Cartesian plane, the distance is in the x-axis and path loss is on the y-axis. By denoting $P_{PL}[dB]$ as y , and $10\log d$ as x, we obtain the linear equation of a line on the origin with the slope of η such as, y = η.x,

where y is the path loss of power between two beacon nodes, x is the logarithm of distance between them and η is the path loss exponent of the environment.

As an example of calculating calibration parameter, let us consider the sector formed by two neighbor beacon nodes B and C like Figure 1. We also have two data points; P1(x1, y1) and P2(x2, y2) where x1 is the distance between beacon node A and beacon node B, y1 is the corresponding pass loss, and x2 is the distance between beacon node A and beacon node C, y2 is the corresponding pass loss respectively. In the Fig. 2, two points P_1 and P_2 are depicted.

Here, we set the objective to find the fitting line with the condition that the sum of the squares of the distances of those points to this line will be minimized. And the main goal is to find a slope η of this line that closely describes this relation.

The sum of squares is expressed as,

$$\omega = \left(y_1 - \eta x_1\right)^2 + \left(y_2 - \eta x_2\right)^2 \tag{8}$$

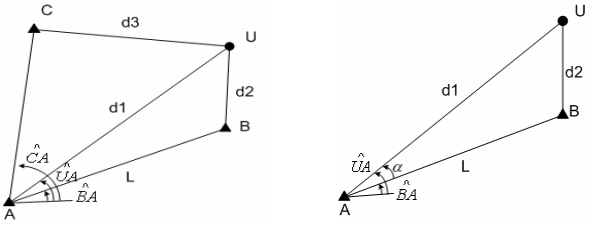
Differentiating the ω to η and setting equal to zero, we can find the slope of the line,

$$\eta = \frac{x_1 y_1 + x_2 y_2}{x_1^2 + x_2^2} \tag{9}$$

The value of path loss exponent according the angle can be periodically stored in beacon node A. At this situation, when an unknown node wants to localize its position, it transmits the signal requesting the angle-calibration parameter to beacon nodes in its range. In the next sub-section, we describe the process in which beacon nodes find the unknown node's direction of arrival and lastly transmit the angle calibration parameter to that unknown node.

3.4 Calculating the Direction of Arrival from Unknown Nodes to Beacon Nodes

In this subsection, the procedure to calculate the relative direction of arrival from unknown node to beacon nodes is illustrated.



(a) (b)

Fig. 3. Directions of arrival from one unknown node to beacon nodes

As showed in Fig. 3.(a), an unknown node U is in the sector formed by beacon nodes A, B and C. When U transmits the signal, the beacon nodes in its transmission range use RSS and path loss exponent to determine the distance from this unknown

node to them. In this case the path loss exponent is the omni-calibration parameter determined by all the signals from their neighbor beacon nodes. After that these distance value will be transmitted to neighbor beacon nodes surrounding. When beacon node A receives more than two distance value from its neighbor beacon nodes, it chooses two sets of date from two beacon nodes that are not adjacent. In Fig.3.(a), beacon node A uses data from beacon node B and C. By using the geometry relation, we can define the direction from unknown node U to beacon node A.

By applying the law of cosines in the triangle formed by two beacon nodes and one unknown node, as showed in Fig. 3.(b), where d1 is the estimated distance from the unknown node to beacon node A and d2 is the estimated distance from the unknown node to beacon node B. We can obtain the angle α such that,

$$\alpha = \cos^{-1}\left(\frac{d1^2 + L^2 - d2^2}{2Ld1}\right) \tag{10}$$

From the value of α, the direction from U to beacon node A is described,

$$\hat{UA} = \left(\hat{BA} \pm \alpha\right)(mod\, 2\pi) \tag{11}$$

There exist two values of $\hat{UA}$ if we just use only one data from beacon node B or C. That is the reason why a beacon node has to know the distances from the unknown node to at least two beacon nodes. By using the intersection between two sets of results, we can find the unique value of direction of arrival.

When a beacon node knows the direction of arrival from an unknown node, a packet containing the unknown node's ID and the path loss exponent in the sector where this unknown node is placed, is transmitted. When this unknown node receives this packet, the path loss exponent is exploited to get the best distance measurement. Ultimately, the lateration is used to determine the unknown location.

3.5 Location Detection

The location of unknown node can be found by the lateration method. This method requires at least 3 beacon nodes that are not in a straight line to compute the position. Figure 4 shows an illustration of lateration method.

Generally, suppose the location of unknown node is denoted by (x, y), the distance from that unknown node to beacon nodes are expressed by,

$$d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{12}$$

Where i is from 1 to N.

Squaring both sides of formula (13) and then subtracting the equation i = 1, yields the following set of equation:

$$2x(x_1 - x_i) + 2y(y_1 - y_i) = x_1^2 + y_1^2 - x_i^2 - y_i^2 - d_1^2 + d_i^2 \tag{13}$$

where i is from 2 to N.

Formula (14) can be expressed in matrix form as,

$$A.\theta = B \tag{14}$$

and

$$A = \begin{bmatrix} (x_1 - x_2) & (y_1 - y_2) \\ (x_1 - x_3) & (y_1 - y_3) \\ ..... & \\ (x_1 - x_N) & (y_1 - y_N) \end{bmatrix}, \quad B = \frac{1}{2}\begin{bmatrix} x_1^2 + y_1^2 - x_2^2 - y_2^2 - d_1^2 + d_2^2 \\ x_1^2 + y_1^2 - x_3^2 - y_3^2 - d_1^2 + d_3^2 \\ ..... \\ x_1^2 + y_1^2 - x_N^2 - y_N^2 - d_1^2 + d_N^2 \end{bmatrix}, \theta = \begin{bmatrix} x \\ y \end{bmatrix},$$

To get the best estimation of location of the unknown node, the standard least squares (LS) method is applied.

Suppose ε is the differentiation between the practical value B and the estimate value in the model $A\hat{\theta}$, the LS method is designed to minimize object function Ψ:

$$\Psi = \xi^T.\xi = (B - A.\theta)(B - A.\theta) \tag{15}$$

The best solution for (16) is given by,

$$\hat{\theta}_{LS} = \left(A^T A\right)^{-1} A^T B \tag{16}$$

The proposed algorithm uses the calibration according to angle to diminish the affects of factors in the communication environment. This is a simple solution but it can get the best estimation of distance from unknown node to beacon node. This method reduces the error in distance measurement and then we apply the least square to get the best estimation of location of unknown node. The proposed algorithm is a simple algorithm and has small computation.

4 Simulation Experiment

To evaluate our algorithm, we use Matlab to simulate the algorithm and further to compare the performance of the proposed algorithm with those of the omni-calibration (self-calibration [5]) and with the case without calibration. The coordinate of beacon nodes and unknown nodes are chose on an area of 100m by 100m. We use 10 unknown nodes for averaging the location error, the unknown nodes' real position are u1(18, 10)m, u2(17, 40)m, u3(27, 30)m, u4(42, 60)m, u5(60, 40)m, u6(62, 64)m, u7(63, 79)m, u8(72, 26)m, u9(78, 50)m, u10(80, 35)m. And there are 7 beacon nodes located in positions: b1(0,0)m, b2(80,0)m, b3(95,45)m, b4(60,70)m, b5(0,80)m, b6(45,10)m, b7(80,80)m. In all simulation scenarios, we just use 5 first beacon nodes except for the final case location errors changing with numbers of beacon nodes. To get the more precise results, we take 20 samples and 50 independent times for every case.

4.1 Simulation Model

To perform the localization algorithm, we create an environment model with the path-loss exponent being changed with the angle around one node. We divide the space around a node into sector and set ones with different values of path loss exponent. We create these values with mean value being fixed at 4.5 and they are randomly obtained around the mean. The maximum declination between them and the mean is characterized by parameter called delta. The value of delta shows the changing level of environment surrounding where the WSN works. After that the white Gaussian noise is added to the path loss of power to generalize the affections in the communication environment.

4.2 Ranging the Distance with Numbers of Independent Measuring

In this part of simulation, we suppose that the variance of white Gaussian noise process is 1dB, and delta is 0.2. The main goal of this part is to show the error in measuring the distance between one anchor and the unknown node u1. In the Fig. 4, x-axis is the number of times calculating independently, while y-axis show the value of distance.

The simulation result shows that the accuracy of distance measuring with our angle -calibration is better than those with omni calibration and non calibration. The times is the number of calculation independent and in each times, the value of path-loss exponent according to each angle will be get randomly in the setting value. When the transition is time to time, our algorithm makes better calibration such that it has better range accuracy than others.

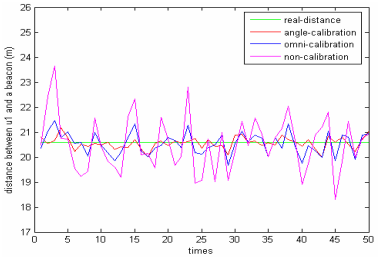


Fig. 4. Distance error according to the number of independent measuring

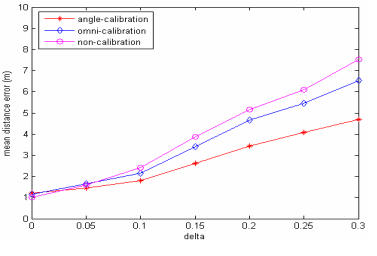


Fig. 5. Mean error of measured distance according to real distance from u1 to beacon nodes

4.3 Mean Error of Distance from u1 to Beacon Nodes with Different Delta

We suppose the variance of white noise is 1 dB, and the value of delta is changing from 0 to 0.3. Fig.5 shows the mean distance error that is average value of difference between the measured distance and real distance from all distance from an unknown node to other beacon nodes in its transmission range. All of these ranging distances will be used in lateration method.

From Fig. 5, we observe that our algorithm has better ranging error than omni-calibration and without case. When delta is small the main factor causing error in system is white Gaussian noise process, both three method have the similar performances in ranging. Clearly, when delta is 0 and 0.05, the mean distance error for angle calibration is 1.20m, 1.43m, comparing with 1.16m, 1.66m in the omni calibration method, and 1.00m, 1.60m in the without case. However, at delta higher, our algorithm is the better solution in ranging than others. As showed in the Fig. 5, with delta is 0.1, 0.2, 0.3, the angle calibration proves a good performance in ranging with mean errors 1.79m, 3.43m, 4.69m respectively; while the omni calibration has higher errors 2.14m, 4.65m, 6.52m, correspondingly according with 2.42m, 5.15m, and 7.53m in the case non-calibration.

4.4 Localize the Unknown Nodes with Delta

In this part of simulation, we also suppose variance of white Gaussian noise is 1 dB. We take this simulation to evaluate the performance of the proposed algorithm versus the omni-calibration and case without. In the Fig. 6, axis is the value of delta characterizing for affections from environment and y is the average error of all estimated positions of unknown nodes comparing with the real unknown node positions.

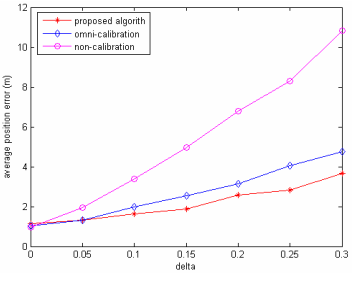


Fig. 6. Relation between location error and delta

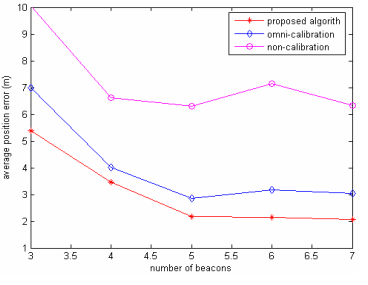


Fig. 7. Average position error according to number of beacon nodes

The results in Fig. 6 show that both three have similar efficiency at small delta, but when delta is increased the algorithms using calibration have better performance. At higher delta, we can find that the location precision of our algorithm is better than that of omni-calibration and without case. Clearly, at delta 0 and 0.05 the average position error of our algorithm are 1.15m and 1.32m comparing with 1.04m and 1.32m with the omni-calibration, whilst the without case obtains the better value 0.95m at delta 0 and worse value 1.94 at delta 0.05. When delta is increased to higher value, the case without calibration has much larger errors; however, our calibration and the omni-calibration have better results; at delta 0.15 our algorithm sustains position error 1.86m comparing with 2.55m in omni-calibration; delta 0.2, the angle-calibration takes position error 2.57m versus 3.15m in omni-calibration; and when delta value is 3.0, our calibration gets 3.66m for location error contrasting to 4.76m in omni-calibration method.

4.5 Relation between Location Errors with Number of Beacon Nodes

We implemented the simulation with the different number of beacon nodes in the area, and in this part of simulation, it is supposed that variance of white Gaussian noise is 1dB. We set value of delta to be 0.2 and record the simulation result with different number of beacon nodes. Fig. 7 demonstrates the simulation result in which x-axis is the number of beacon nodes and y-axis is the average position error value.

From the simulation result, we can observe that the angle calibration has lower location error than omni calibration and without case with the same number of beacon nodes, and the position error will reduce as the number of beacon nodes increases for all cases. More details, with 3 beacon nodes, our algorithm has location error 5.39m, while 6.99m and 10.08m in omni-calibration and without case respectively; from 4 beacon nodes to beyond, our algorithm always has the better results 3.46m, 2.18m, 2.15m, and 2.05m comparing with 4.02m, 2.85m, 3.17m, and 3.03m, respectively in omni-calibration while the non-calibration has large location errors.

5 Conclusion

Wireless sensor networks have numerous applications, and the localization is the fundamental requirement for WSNs applications. RSS is one of the simplest ranging methods for easy and simple but the concession is performance. When a system using RSS associated with angle calibration, the precision is improved more. From the better ranging it has better location detection with LS method. The proposed algorithm takes the merits of two methods. The algorithm is simple, computation is small and it has a higher precision.

Acknowledgement. This work was supported by the second stage of the Brain Korea 21 project in 2008. This work was also supported in part by the Ministry of Commerce, Industry, and Energy and in part by Ulsan Metropolitan City through the Network-based Automation Research Center at the University of Ulsan.

References

1. Cheung, K.W., So, H.C., Ma, W.K., Chan, Y.T.: A Constrained Least Squares Approach to Mobile Positioning: Algorithms and Optimality. EURASIP Journal on Applied Signal Processing 23 (2006)
2. Tian, H., Wang, S., Xie, H.: Localization Using Cooperative AOA Approach. In: The International Conference on Wireless Communications, Networking and Mobile Computing, WiCom 2007, September, pp. 2416–2419 (2007)
3. Cheung, K.W., So, H.C., Ma, W.K., Chan, Y.T.: Received Signal Strength Based Mobile Positioning via Constrained Weighted Least Squares. In: Proceedings of IEEE International Conference on, Acoustics, Speech, and Signal Processing (ICASSP 2003), April, vol. 5 (2003)

4. Elnahrawy, E., Xiaoyan, L., Martin, R.P.: The Limits of Localization Using Signal Strength: A Comparative Study. Sensor and Ad Hoc Communications and Networks. In: First Annual IEEE Communications Society Conference on IEEE SECON 2004, October, pp. 406–414 (2004)
5. Kang, J., Kim, D., Kim, J.: RSS Self-calibration Protocol for WSN Localization. In: 2nd International Symposium on Wireless Pervasive Computing, ISWPC 2007 (February 2007)
6. Constatino, P.V., Garcia, J.L.: A Simple Approach to a Statistical Path Loss Model for Indoor Communications. In: 27thEuropean Microwave Conference, October, vol. 1, pp. 617–623 (1997)

A General Framework for Adaptive Filtering in Wireless Sensor Networks Based on Fuzzy Logic*

Hae Young Lee and Tae Ho Cho

School of Information and Communication Engineering, Sungkyunkwan University
Suwon 440-746, Republic of Korea
{sofware,taecho}@ece.skku.ac.kr

Abstract. Wireless sensor networks are vulnerable to false data injection attacks. In these, adversaries inject false reports into the network using compromised nodes, with the goal of deceiving the base station and depleting the energy resources of forwarding nodes. Several filtering schemes, and the adaptive versions of them, have been proposed to detect and drop such false reports during the forwarding process. In this paper, we propose a fuzzy-based framework to achieve adaptive filtering of false reports. A fuzzy rule-based system controls the security level of the filtering scheme through the determination of the required number of endorsements in every report. Compared to existing adaptive solutions, the proposed method uses more practical factors for the determination and can be applied to the network, regardless of scale. The proposed method can provide energy saving and sufficient resilience against false data injection attacks.

Keywords: Wireless sensor networks, false data injection attacks, false data filtering, fuzzy logic, security.

1 Introduction

Due to the many attractive characteristics of sensor nodes, wireless sensor networks (WSNs) have been adapted to applications, including smart homes and laboratories, remote environment monitoring, in-plant robotic control and military surveillance [1]. In many applications, sensor nodes are deployed in open environments, and hence can be physically captured by adversaries [2]. Compromised nodes can be used to launch false data injection attacks [3] in which they report nonexistent events appearing at arbitrary locations in the field [4]. Forged reports will not only cause false alarms that waste real-world response efforts, but also drain the finite amount of energy in a battery-powered network [5]. Researchers first proposed basic filtering schemes [3,6-8] to address false data injection attacks, through symmetric key sharing among sensor nodes. Then, adaptive versions [9-12] of the schemes were proposed to achieve better energy-efficient filtering of false reports, through the determination of the required

* This research was supported by the MKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement). (IITA-2008-C1090-0801-0028).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1234–1241, 2008.
© Springer-Verlag Berlin Heidelberg 2008

number of message authentication codes (MACs) per report. The adaptive filtering schemes can provide energy saving and sufficient resilience against false data injection attacks. However, they have the following limitations: 1) they use the number of compromised nodes in inference. In real-world WSNs, it may be impractical to get the exact number of compromised nodes. 2) They can address only a limited number of cluster nodes (e.g., 21 cluster nodes). Each cluster may consist of a large number of nodes in large-scale WSNs.

This paper presents a general framework for adaptive filtering (FAF) in WSNs. A fuzzy rule-based system is exploited to determine the number of MACs required for energy-efficient filtering of false reports. The rate of dropped reports, the rate of false reports, and the feedback of the fuzzy system are used to determine this. The contributions of FAF are: 1) it does not require the number of compromised nodes in the inference. 2) There is no limit to the number of cluster nodes. 3) It can be applied to various filtering schemes, such as key inheritance-based filtering (KIF) [8].

2 Background

In this section, we briefly describe basic and adaptive filtering schemes.

2.1 Basic Filtering Schemes

The basic filtering schemes, such as statistical en-route filtering (SEF) [3], interleaved hop-by-hop authentication (IHA) [6], dynamic en-route filtering (DEF) [7], and key inheritance-based filtering (KIF) [8], defend against false data injection attacks, through symmetric key sharing among sensor nodes. Each of them provides its own method to share keys among nodes. On the other hand, sensing reports are generated and verified in a similar fashion. When an event (e.g., a tank) occurs, multiple sensing nodes collaboratively generate a sensing report. The report is endorsed by attaching multiple MACs generated by them using their keys. As the report is forwarded towards the base station (BS) over multiple hops, each forwarding node verifies some MACs in the report using its own keys that are shared with some of the sensing nodes. The report is dropped if the verification fails. There is a trade-off between security and overhead in the choice of the required number of MACs per report (RMR) [3]. A large RMR makes forging reports more difficult, but it consumes more energy in forwarding. Conversely, a small RMR may make the filtering scheme inefficient or even useless if adversaries have compromised a large number of nodes [13]. In the basic schemes, the network designer is totally responsible for the choice of RMR.

2.2 Adaptive Filtering Schemes

The adaptive filtering schemes – adaptive SEF (ASEF) [9], AIHA [10], AKIA [11], and ADEF (ADEF) [12] – achieve better energy-efficiency in false data filtering, through the adaptive control of RMR. They exploit fuzzy rule-based systems to determine RMR that can minimize energy consumption without sacrificing security. Thus, they can provide both energy saving and sufficient resilience against false data injection attacks. One of their major drawbacks is that they all use the number of compromised nodes in the inference. In real-world WSNs, it may be impractical to

obtain the exact number of compromised nodes, since adversaries may attempt to suppress such information. Overestimating of the number of compromised nodes can make the filtering scheme energy-inefficient; underestimating may lead the filtering scheme to be totally useless. Another drawback is they can address only a limited number of cluster nodes. In large-scale WSNs, each cluster may comprise a large number of nodes (e.g., several dozen nodes). To adopt adaptive filtering on such networks, its fuzzy rule-based system may require to be wholly redesigned.

3 Fuzzy-Based Framework for Adaptive Filtering

In this section, we describe FAF in detail.

3.1 Assumptions

We consider a WSN composed of a large number of small sensor nodes. One of the basic filtering schemes described in Section 2.1 is employed in the network. We assume that BS cannot be compromised. We also assume that BS can know the rate of dropped reports (DRR) and the rate of reports rejected by BS (RRR). To measure DRR, we can deploy tamper-resistant nodes that only record the number of dropped reports and report such information to BS (Fig. 1(a)) [14]. Each of these nodes may cover several clusters in the network.

3.2 Overview

FAF works on the basic scheme. To achieve energy-efficient filtering of false reports, BS periodically determines whether the required number of MACs per report (RMR) needs to be increased, decreased, or held with a fuzzy rule-based system. DRR, RRR, and the feedback of the fuzzy system (FSF) are used for the inference (Fig. 1(b)). The conclusions (Fig. 1(c)) of the fuzzy system are: 1) hold RMR (H). 2) Increase RMR by one (I1). 3) Increase RMR by two (I2). 4) Decrease RMR by one (D). 5) Turn off the scheme (TO). If the fuzzy system concludes RMR is to be increased or decreased, BS broadcasts a control message to all corresponding nodes (Fig. 1(d)). The filtering scheme is turned off when the fuzzy system outputs TO.

3.3 Factors for the Determining

In FAF, DRR, RRR, and FSF are used for the fuzzy inference. Through DRR and RRR, we can determine if RMR needs to be increased. Basically, high DRR indicates that a large number of false reports has been injected into the network, but detected by the filtering scheme. That is, the network comes under attack, but is protected by the filtering scheme. Thus, RMR may not need to be increased. On the other hand, very low DRR implies that the network is not under false data injection attacks, or the filtering scheme cannot detect any false reports. We also may not need to increase RMR in the former case. However, in the latter case, adversaries may have sufficient compromised nodes to generate correct MACs for false reports. A large RMR makes forging reports more difficult [3]. Therefore, RMR should be increased until the scheme once again can provide sufficient resilience. RRR can be used to determine if

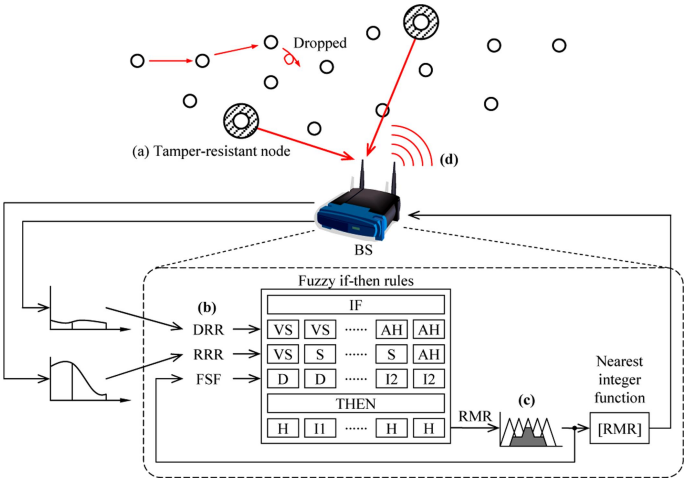


Fig. 1. FAF overview

the filtering scheme works well. Very low RRR indicates that the network is under the protection of the scheme. Thus, RMR may not need to be increased. Conversely, high RRR implies that the filtering scheme does nothing but consume energy. Therefore, RMR should need to be increased.

FSF can be used to stop infinitely increasing RMR. A large RMR consumes more energy in the forwarding. Thus, if RMR has been increased based on the prior conclusion, we may as well suppress the increment of RMR. FSF can be also used to determine if we need to turn off the filtering scheme. If adversaries have compromised a very large number of nodes, the filtering scheme may not filter out any false reports generated by the adversaries. Under this situation, we may as well turn off the filtering scheme to eliminate the overhead for MAC generation. If RRR is still high after the quiet increment of RMR, it may be a time to turn off the filtering scheme. Adversaries may possess sufficient compromised nodes to disable the filtering scheme. Therefore, we should make a decision based on FSF.

3.4 Fuzzy-Based Determining

FAF exploits a fuzzy rule-based system to control RMR. Part of the appeal of fuzzy rule-based systems is that they can be used for approximate reasoning. This is particularly important when there is uncertainty in the reasoning processes, in addition to imprecision in data [15]. In the reasoning, DRR and FRR are the primary sources of uncertainty. In real-world WSNs, legitimate reports may be dropped during the forwarding process, since nodes may misbehave [16]. For to the same reason, uncompromised nodes may generate incorrect MACs for real events. Thus, legitimate reports may be rejected by BS. Therefore, it is necessary to use approximate reasoning to handle such fuzzy information.

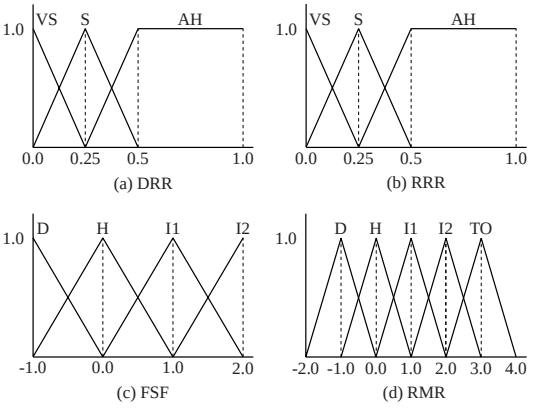


Fig. 2. The membership functions of the fuzzy rule-based system

3.5 Fuzzy Rule-Based System

The membership functions of the three input parameters – (a) DRR, (b) RRR, and (c) FSF – of the fuzzy rule-based system are shown in Fig. 2. The labels in the fuzzy variables are presented as follows.

- DRR/RRR = {VS (Very Small), L (Low), AH (Above Half)}.
- FSF = {D (Decreased), H (Held), I1 (Increased by 1), I2 (Increased by 2)}.

The membership functions of the output parameter of the fuzzy system are illustrated in Fig. 2(d). The labels in the fuzzy variables are presented as follows:

- RMR = {D, H, I1, I2, TO}.

Table 1 shows some rules of the fuzzy rule-based system. The fuzzy system basically prefers to increase rather than decrease RMR, since more nodes are compromised as time elapses [17]. If rejected reports are a small portion of the total, the network is not under threat or most false reports are detected by the filtering scheme during the forwarding process. Thus, the fuzzy system holds RMR (Rule 10). A high RRR indicates that the filtering scheme cannot drop most false reports due to a large number of compromised nodes. Thus, the fuzzy system increases RMR to restore the resilience against the false data injection attacks. The increment rate is determined based on the prior conclusion (Rule 12). If RMR has been quietly increased according to the prior conclusion, the fuzzy system regards the resilience of the scheme as unrecoverable. Therefore, the fuzzy system turns off the filtering scheme to conserve energy consumed by computations (Rule 30). RMR is decreased only when RMR has quietly increased and very low FRR and RRR are achieved as the result of such increment (Rule 28).

Table 1. Fuzzy if-then rules

Rule No.	DRR	RRR	FSF	RMR
10	VS	VS	H	H
12	VS	AH	H	I2
28	VS	VS	I2	D
30	VS	AH	I2	TO

4 Simulation Results

To demonstrate the effectiveness of FAF, we compare FAF with the basic filtering schemes and the adaptive filtering schemes through simulation. The parameters used in the simulation are equivalent to those of [9-10].

Fig. 3 shows the average energy consumption per report delivery in the IHA-based network under a practical assumption in which BS overestimates or underestimates the number of compromised nodes. The number of compromised nodes is between 0 and 20. FAF (filled circles) on IHA is the most energy-efficient, since it does not require the number of compromised nodes in the inference. Conversely, in AIHA, the overestimation (empty rectangles) of the number of compromised nodes may lead to the choice of a large RMR, and thus consume more energy in forwarding. Moreover, AIHA may turn the filtering capability off, even if adversaries have compromised a small number of nodes. In the figure, AIHA turned the filtering capability off when there were seventeen compromised nodes, even if the network is resilient for up to 20 compromised nodes. Thus, any false reports cannot be filtered after the turnoff. Underestimation (empty diamonds) of the number of compromised nodes may result in the choice of a small RMR, less than the number of compromised nodes. Therefore, AIHA cannot detect false reports during the forwarding process. That is, AIHA does nothing but consume extra energy. Therefore, the use of AIHA for the false data filtering is not encouraged if it is very difficult or impossible to get the exact number of compromised nodes.

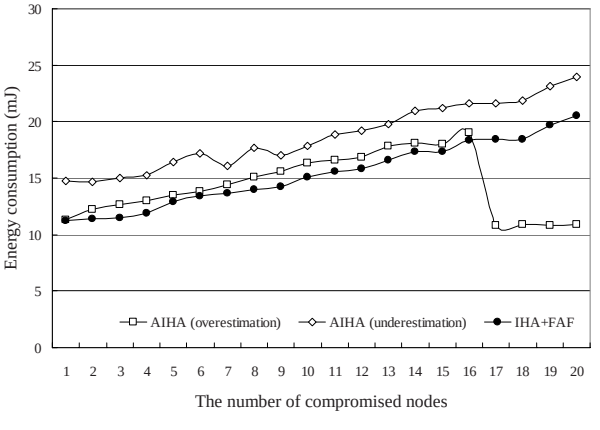


Fig. 3. The energy consumption per report delivery in the IHA-based network

Fig. 4 shows the average energy consumption per report delivery in the SEF-based network. We assumed that BS overestimates the number of compromised nodes. A very small RMR drastically decreases the detection power of SEF, and thus can make false reports undetected during the forwarding process. Conversely, a large RMR consumes more energy in forwarding, but can lead to the early detection of false reports. Thus, the underestimation of the number of compromised nodes in ASEF may cause the worst result. Conversely, as shown in the figure, the overestimation (empty rectangles) can lead to energy saving, if a small number of nodes are compromised. However, if the number of compromised nodes exceeds seventeen, it cannot detect false reports and thus consumes more energy. FAF on SEF (filled circles) is less energy-efficient with a small number of compromised nodes. It is more efficient in terms of energy saving than ASEF and can detect false reports when adversaries have compromised a large number of nodes.

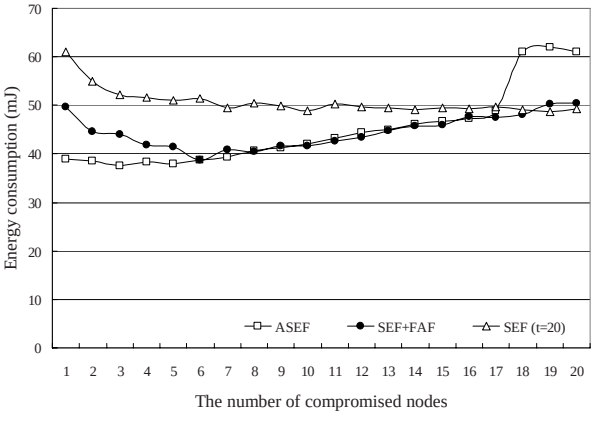


Fig. 4. The energy consumption per report delivery in the SEF-based network

5 Conclusion and Future Work

In this paper, we proposed FAF for WSNs. A fuzzy rule-based system is exploited to determine if RMR needs to be varied for energy-efficient filtering of false reports with the consideration of DRR, RRR, and FSF. FAF can be applied to various filtering schemes, regardless of network scale, and provide both energy saving and sufficient resilience. We plan to simulate for FAF on other filtering schemes. Finally, the specific optimization of FAF for SEF should be examined.

References

1. Suzuki, R., Sezaki, K., Tobe, Y.: A Protocaol for Policy-Based Session Control in Disruption Tolerant Sensor Networks. IEICE Trans. Commun. E90-B(12), 3426–3433 (2007)
2. Przydatek, B., Song, D., Perrig, A.: SIA: Secure Information Aggregation in Sensor Networks. In: The ACM Conference on Embedded Networked Sensor Systems, pp. 255–265 (2003)

3. Ye, F., Luo, H., Lu, S.: Statistical En-Route Filtering of Injected False Data in Sensor Networks. IEEE J. Sel. Area Comm. 23(4), 839–850 (2005)
4. Yang, H., Lu, S.: Commutative Cipher Based En-Route Filtering in Wireless Sensor Networks. In: The IEEE Vehicular Technology Conference, pp. 1223–1227 (2003)
5. Li, F., Wu, J.: A Probabilistic Voting-Based Filtering Scheme in Wireless Sensor Networks. In: The International Wireless Communications and Mobile Computing Conference, pp. 27–32 (2006)
6. Zhu, S., Setia, S., Jajodia, S., Ning, P.: An Interleaved Hop-by-Hop Authentication Scheme for Filtering of Injected False Data in Sensor Networks. In: The IEEE Symposium on Security and Privacy, pp. 259–271 (2004)
7. Yu, Z., Guan, Y.: A Dynamic En-Route Scheme for Filtering False Data Injection in Wireless Sensor Networks. In: The IEEE International Conference on Computer Communications, pp. 1–12 (2006)
8. Lee, H.Y., Cho, T.H.: Key Inheritance-Based False Data Filtering Scheme in Wireless Sensor Networks. In: Madria, S.K., Claypool, K.T., Kannan, R., Uppuluri, P., Gore, M.M. (eds.) ICDCIT 2006. LNCS, vol. 4317, pp. 116–127. Springer, Heidelberg (2006)
9. Lee, H.Y., Cho, T.H.: Fuzzy Based Security Threshold Determining for the Statistical En-Route Filtering in Sensor Networks. Enformatika 14, 157–160 (2006)
10. Lee, H.Y., Cho, T.H.: Fuzzy-Based Adaptive Threshold Determining Method for the Interleaved Authentication in Sensor Networks. In: Gelbukh, A., Reyes-Garcia, C.A. (eds.) MICAI 2006. LNCS (LNAI), vol. 4293, pp. 112–121. Springer, Heidelberg (2006)
11. Lee, H.Y., Cho, T.H.: Fuzzy Adaptive Threshold Determining in the Key Inheritance Based Sensor Networks. In: Okuno, H.G., Ali, M. (eds.) IEA/AIE 2007. LNCS (LNAI), vol. 4570, pp. 64–73. Springer, Heidelberg (2007)
12. Lee, S.J., Lee, H.Y., Cho, T.H.: A Threshold Determining Method for the Dynamic Filtering in Wireless Sensor Networks Based on Fuzzy Logic. International Journal of Computer Science and Network Security 8(4) (2008)
13. Zhang, W., Cao, G.: Group Rekeying for Filtering False Data in Sensor Networks: A Predistribution and Local Collaboration-based Approach. In: The IEEE International Conference on Computer Communications, pp. 503–514 (2005)
14. Lee, H.Y., Cho, T.H.: Fuzzy Adaptive Selection of Filtering Schemes for Energy Saving in Sensor Networks. IEICE Trans. Commun. E90-B(12), 3346–3353 (2007)
15. Serrano, N., Seraji, H.: Landing Site Selection using Fuzzy Rule-Based Reasoning. In: The IEEE International Conference on Robotics and Automation, pp. 4899–4904 (2007)
16. Marti, S., Giuli, T.J., Lai, K., Baker, M.: Mitigating Routing Misbehavior in Mobile Ad hoc Networks. In: ACM International Conference on Mobile Computing and Networking, pp. 255–265 (2000)
17. De, P., Liu, Y., Das, S.K.: Modeling Node Compromise Spread in Wireless Sensor Networks Using Epidemic Theory. In: International Workshop on Wireless Mobile Multimedia, pp. 237–243 (2006)

Stability Analysis of AQM Algorithm Based on Generalized Predictive Control*

Xing-hui Zhang[1], Kuan-sheng Zou[1], Zeng-qiang Chen[2], and Zhi-dong Deng[3]

[1] Department of computer science, Tianjin University of Technology and Education, Tianjin, 300222, China
[2] Department of Automation, Nankai University, Tianjin, 30071, China
[3] State Key Laboratory of Intelligence Technology and Systems, Tsinghua University, Beijing, 10084, China
xhzhang@tute.edu.cn

Abstract. The network congestion has become increasingly serious. So many active queue management (AQM) algorithms used in routers have been proposed to alleviate network congestion. Generalized predictive control (GPC) is an effective control strategy, which has been widely applied in nonlinear and time-delay systems. This paper mainly analyses the stability of GPC as an AQM algorithm. It educes the discrete model of the network based on the fluid-flow model of TCP congestion window, then elicits the closed loop transfer function of router's output queue length with reference queue length. When the length of predictive region isn't smaller than system total time delay, it gives the characteristic equation of the closed loop system in one-step control. Then conditions of system stabilization are proved with jury criterion, and it gives three theorems about system stability. The validity of theorems is verified by simulation results.

Keywords: active queue management (AQM), stability analysis, network congestion, generalized predictive control (GPC), round trip time.

1 Introduction

With the increasing application of the multimedia service in the internet, the network congestion has become increasingly serious. So many active queue management (AQM) algorithms used in routers have been proposed to alleviate network congestion. AQM algorithm drops packets before buffer becomes full and use average queue length as an congestion indicator. Drop tail algorithm is the first one to solve, it is simple to implement but has some shortages. Literature [1] proposed RED algorithm, but its performance was closely related with parameters configuration. Imprecise parameters configuration would lead to the greatly fluctuate of queue length. Literature [2] presented an TCP flow control model with AQM algorithm, and literature [3] gave linearization

* This work is supported by 863 High technology Research and Development Project (2006AA04Z208) and Tianjin Natural Science Foundation Projects (06YFJMJC01600, 08JCZDJC21900).

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1242–1249, 2008.
© Springer-Verlag Berlin Heidelberg 2008

model of this TCP flow control model and presented proportion – Integral (PI) AQM control strategy. PI algorithm has smaller queue oscillation than RED algorithm, but the parameters usually come from men's trial, so the performance is not guaranteed to be optimal. The literature [4] presented a PID algorithm with stability region, it can get better system transient performance than PI algorithm, but its parameters setting are rather complicated.

Generallized predictive control (GPC) [5] is an effective control strategy, which has been widely applied in nonlinear and time-delay systems. Improved of GPC algorithm [6] and its robustness and stability analysis have become a hotspot [7]. Literature [8] used fast GPC as an AQM algorithm in network congestion control, and achieved a relatively good control effect, but stability analysis of this algorithm is not given. This paper mainly analyses the stability of GPC as an AQM algorithm. Firstly, it educes the discrete model of the network based on the fluid-flow model of TCP congestion window, then it elicits the closed loop transfer function of router's output queue length with reference queue length. Secondly, when the length of predictive region isn't smaller than system total time delay, it gives the characteristic equation of the closed loop system in one-step control. Then conditions of system stabilization are proved with jury criterion, and it gives three theorems about system stability. The validity of theorems is verified by simulation results.

2 TCP Flow Control Model

Misra and others established dynamic model of congestion window under the AQM / TCP connection in literature [2]. It is described as equation (1).

$$\begin{cases} \dot{w}(t) = \dfrac{1}{R(t)} - \dfrac{w(t)\omega(t-R(t))}{2R(t-R(t))}\, p(t-R(t)) \\ \dot{q}(t) = \dfrac{w(t)}{R(t)} N(t) - C \end{cases} \tag{1}$$

Here w is the TCP window size; q is queue length; R is round trip time; C is bottleneck capacity; N is load factor; p is packet dropping probability. In order to facilitate the analysis of system stability, Hollot gave linearization model based on small signal theory as equation (2).

$$\begin{cases} \delta \dot{w}(t) = -\dfrac{2N}{R_0^2 C}\delta w(t) - \dfrac{R_0 C^2}{2N^2}\delta p(t-R_0) \\ \delta \dot{q}(t) = \dfrac{N}{R_0}\delta w(t) - \dfrac{1}{R_0}\delta q(t) \end{cases} \tag{2}$$

Assume that $\dfrac{\delta q(s)}{\delta p(s)} = P(s)$, it gets equation (3) through Laplace's transformation of equation (2).

$$P(s) = \frac{(R_0 C)^3 / (2N)^2 e^{-R_0 s}}{(R_0^2 C s / 2N + 1)(R_0 s + 1)} \tag{3}$$

Translate equation (3) to general pulse transfer function with zero-order holder, which is shown as equation (4). Here the sample time is T.

$$P(z^{-1}) = K\frac{(b_d + b_{d+1}z^{-1})z^{-(d+1)}}{1 + a_1 z^{-1} + a_2 z^{-2}} \tag{4}$$

The coefficients value of equation (4) is given in equation (5). Where $T_1 = R_0$, $T_2 = \frac{R_0^2 C}{2N}$ ($T_1 \neq T_2$), $K = \frac{(R_0 C)^3}{(2N)^2 (T_1 - T_2)}$ $d = \left\lfloor \frac{R_0}{T} \right\rfloor$, ($\left\lfloor \frac{R_0}{T} \right\rfloor$ means the smallest integer that bigger than $\frac{R_0}{T}$).

$$\begin{aligned} a_1 &= -(e^{-\frac{T}{T_2}} + e^{-\frac{T}{T_1}}), \quad b_d = (T_1 + T_2) - (T_2 e^{-\frac{T}{T_2}} + T_1 e^{-\frac{T}{T_1}}) \\ a_2 &= e^{-(\frac{1}{T_2} + \frac{1}{T_1})T}, \quad b_{d+1} = (T_1 + T_2)e^{-(\frac{1}{T_2} + \frac{1}{T_1})T} - (T_2 e^{-\frac{T}{T_1}} + T_1 e^{-\frac{T}{T_2}}) \end{aligned} \tag{5}$$

3 GPC Control Strategy

GPC algorithm based on CARIMA model is given in equation (6).

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + \varepsilon(t)/\Delta \tag{6}$$

Let output value $y(t) = q(t)$, and control value $u(t) = p(t)$. $\varepsilon(t)$ is random noise of network. It shows that the discrete TCP flow control model has the form of the CARIMA model, so GPC algorithm can be used as AQM control strategy. Contrast equation (4) with equation (6), then it gets equation (7).

$$A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2}, \quad B(z^{-1}) = (b_d + b_{d+1}z^{-1})z^{-d} \tag{7}$$

Control objective function is defined as equation (8). Here N_y is predictive domain, *Nu* is control domain, λ is weighted factor, $\lambda \geq 0$.

$$J = E\{\sum_{j=1}^{N_y}[y_r(t) - y(t+j)]^2 + \lambda\sum_{j=1}^{N_u}[\Delta u(t+j+1)]^2\} \tag{8}$$

Obtaining the optimal output value that *j* step forward, it must solve the following diophsotine equations.

$$\begin{cases} 1 = A(z^{-1})\Delta E_j(z^{-1}) + z^{-j}F_j(z^{-1}) \\ E_j(z^{-1})B(z^{-1}) = G_j(z^{-1}) + z^{-j}H_j(z^{-1}) \end{cases} \tag{9}$$

Where $E_j(z^{-1}) = e_0 + e_1 z^{-1} + \cdots + e_{j-1} z^{-j+1}$, $F_j(z^{-1}) = f_0 + f_1 z^{-1} + \cdots + f_{j-1} z^{-j+1}$ $H_j(z^{-1}) = h_0 + h_1 z^{-1} + \cdots + h_{j-1} z^{-j+1}$,

$$G = \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & g_0 \\ g_{N-1} & \cdots & \cdots & g_{N-N_u} \end{bmatrix}_{N_y \times N_u}$$

Minimize equation (8), the GPC law can be get as equation (10) shows.

$$\Delta u(t) = h_c^T [y_r(t) - \bar{F}(z^{-1}) y(t) - z^{-1} \bar{H}(z^{-1}) \Delta u(t)] \tag{10}$$

Where

$$\begin{aligned} &\bar{F}(z^{-1})^T = [F_1(z^{-1}), \cdots, F_{N_y}(z^{-1})] \quad \bar{H}(z^{-1})^T = [H_1(z^{-1}), \cdots, H_{N_y}(z^{-1})] \\ &h_c^T = [1, 0, \cdots, 0](\lambda I + G^T G)^{-1} G^T = \left[\bar{h}_1, \bar{h}_2, \cdots \bar{h}_{N_y} \right] \end{aligned} \tag{11}$$

4 Stability Analysis of Closed-Loop System

It sets equation (12) to simplify the following calculation.

$$T(z^{-1}) = 1 + z^{-1} h_c^T \bar{H} \quad , \quad S(z^{-1}) = h_c^T \bar{F} \tag{12}$$

From equation (10-12), The closed-loop transfer function with output queue length y to the reference queue length y_r is obtained in equation (13). The characteristic equation of closed-loop system is given in equation (14).

$$\frac{y(t)}{y_r(t)} = \frac{z^{-1} B(z^{-1}) h_c^T}{A(z^{-1}) T(z^{-1}) \Delta + z^{-1)} B(z^{-1}) S(z^{-1})} \tag{13}$$

$$Q(z^{-1}) = A(z^{-1}) T(z^{-1}) \Delta + z^{-1)} B(z^{-1}) S(z^{-1}) \tag{14}$$

Lemma 1[9]. The first j polynomial coefficients of $G_j(z^{-1})$ is the object's first *j-th* sample values with step response, which is defined as follows in the convergence region.

$$\frac{z^{-1} B(z^{-1})}{A(z^{-1}) \Delta} = g_0 z^{-1} + g_1 z^{-2} + \cdots + g_j z^{-j+1} + \cdots$$

Theorem 1. The coefficients of $A(z^{-1})$, $B(z^{-1})$ and $G_j(z^{-1})$ meet the following relations.

$$\begin{cases} g_d = b_d \\ (g_{d+1} - g_d) + g_d a_1 = b_{d+1} \\ \vdots \\ (g_{nb} - g_{nb-1}) + (g_{nb-1} - g_{nb-2})a_1 + \cdots + (g_{nb-na} - g_{nb-na-1})a_{nb} = b_{nb} \\ (g_i - g_{i-1}) + (g_{i-1} - g_{i-2})a_1 + \cdots + (g_{nb-na} - g_{nb-na-1})a_{nb} = 0 \qquad (i \geq nb+1) \end{cases}$$

Where $A(z^{-1}) = 1 + a_1 z^{-1} + \cdots a_{na} z^{-na-1}$, $B(z^{-1}) = b_0 + b_1 z^{-1} + \cdots b_{nb} z^{-nb}$

Proof: From Lemma 1; it gets the form of theorem 1.

Lemma 2 [10]. Characteristic polynomial of GPC closed-loop system is expressed as follows:

$$Q(z^{-1}) = A\Delta + z^{-1} h_c^T \bar{S}\ , \quad \text{Where}$$

$$\bar{S} = [z(B - A\Delta G_1), z^2(B - A\Delta G_2), \cdots z^{N_y}(B - A\Delta G_{N_y})]^T$$

Connect equation (9), (12), and (14), then lemma 2 can be educed.

Theorem 2. The characteristic polynomial of closed-loop system (16) is:

$$Q(z^{-1}) = 1 + (a_1 - 1 + \frac{v}{m})z^{-1} + \frac{\lambda}{m}(a_2 - a_1)z^{-2} - \frac{\lambda}{m} a_2 z^{-3},$$

$$\text{Where } m = \lambda + \sum_{j=d}^{N_y - 1} g_j^2 \,, v = \sum_{j=d}^{N_y - 1} g_j g_{j+1}$$

Proof: let $\bar{S}(z^{-1}) = [\bar{S}_1, \bar{S}_2, \cdots \bar{S}_{N_y}]^T$, from Lemma 1, it can get:

$g_0 = g_1 = \cdots g_{d-1} = 0$, $g_d = b_d$. so $\bar{S}(z^{-1}) = 0 \quad (j \leq d)$. When $j \geq d+1$, according to Lemma 2 and theorem 1, $\bar{S}(z^{-1})$ can educed as follows:

$$\bar{S}_{d+2} = g_{d+2} + [g_{d+1}(a_1 - a_2) + g_d a_2]z^{-1} + g_{d+1} a_2 z^{-2}$$

$$\begin{aligned} \bar{S}_j &= z^{j-d}\{(b_d + b_{d+1}z^{-1}) - (1 + a_1 z^{-1} + a_2 z^{-2})[g_d + (g_{d+1} - g_d)z^{-1} \\ &+ \quad \cdots + (g_{j-1} - g_{j-2})z^{-j+d+1} - g_{j-1} z^{-j+d}]\} \\ &= (b_d - g_d)z^{j-d} + [b_{d+1} - g_d a_1 + (g_{d+1} - g_d)]z^{j-d-1} \\ &- [g_d a_2 + a_1(g_{d+1} - g_d) + (g_{d+2} - g_{d+1})]z^{j-d-2} + \cdots \\ &+ [g_{j-1} - a_1(g_{j-1} - g_{j-2}) - a_2(g_{j-2} - g_{j-3})] \\ &+ [g_{j-1} a_1 - (g_{j-1} - g_{j-2})a_2]z^{-1} + g_{j-1} a_2 z^{-2}\} \\ &= g_j + [g_{j-1}(a_1 - a_2) + g_{j-2} a_2]z^{-1} + g_{j-1} a_2 z^{-2} \qquad (j \geq d+3) \end{aligned}$$

Connect these equations with lemma 2, it can get theorem 2.

Theorem 3. When system guarantees condition **H.a** and **H.b**, system is stability.

$\sum_{j=d}^{N_{y-1}} g_j(1-a_2 g_{j+1})$ and $b_d + b_{d+1}$ have the same sign. H.a

$$\lambda > \max\{0, \frac{(a_1-2)\sum_{j=d}^{N_y-1} g_j^2 + \frac{1}{a_2}\sum_{j=d}^{N_{y-1}} g_j}{2(1-a_1+a_2)}, \overline{\lambda}\} \quad \text{H.b}$$

Proof: according to theorem 2 and Jury criterion, the system stability needs the following formulas in equation (15) be founded.

$$\begin{gathered} |Q_3| < 1,\ 1+Q_1+Q_2+Q_3 > 0 \\ -1+Q_1-Q_2+Q_3 < 0,\ \begin{vmatrix} 1 & Q_3 \\ Q_3 & 1 \end{vmatrix} < \begin{vmatrix} Q_2 & Q_1 \\ Q_3 & 1 \end{vmatrix} \end{gathered} \quad (15)$$

Assume $l = \sum_{j=d}^{N_y-1} g_j^2$, then $m = \lambda + l$,so it can get equation (16).

$$1+Q_1+Q_2+Q_3 = a_1 + \frac{v}{m} + \frac{\lambda}{m}(a_2-a_1) - \frac{\lambda}{m}a_2 = a_1\frac{l}{m} + \frac{v}{m} > 0 \quad (16)$$

Then $(b_d + b_{d+1})\sum_{j=d}^{N_{y-1}} g_j(1-a_2 g_{j+1}) > 0$, when **H.a** is guaranteed. Because

$$-1+Q_1-Q_2+Q_3 = -2 + a_1(1+\frac{\lambda}{m}) + \frac{v}{m} - \frac{2\lambda}{m}a_2 < 0 \quad (17)$$

It can be simplified as follows

$$\lambda > \frac{(a_1-2)\sum_{j=d}^{N_y-1} g_j^2 + \frac{1}{a_2}\sum_{j=d}^{N_{y-1}} g_j}{2(1-a_1+a_2)} \quad (18)$$

Q_3 can be simplified in equation (19).

$$Q_3 = \left|\frac{\lambda}{m}a_2\right| < \left|\frac{\lambda}{m}\right| < 1 \quad (19)$$

Simplify equation (20-d), when $\lambda > \overline{\lambda}$, system is stability. Where

$$\bar{\lambda} = \frac{a_2 v - (2 + a_1 - a_2 a_1) l + \sqrt{[(2 + a_1 - a_2 a_1) l - a_2 v]^2 - 4v^2 (1 + a_1 - a_2^2 - a_1 a_2)}}{2(1 - a_2)(1 + a_2 + a_1)} \tag{20}$$

Where $[(2 + a_1 - a_2 a_1) l - a_2 v]^2 - 4v^2 (1 + a_1 - a_2^2 - a_1 a_2) > 0$

Theorem 3 can be proved from equation (15~20).

5 Simulation Results

The simulation is done using MATLAB 7.0 on a PC whose main frequency is 3.0GHz. The network parameters are N=60, R=0.25s, C=3750 packege/s, $T = 0.0675$ s. So the time delay d = 4. According to the theorem, it sets the reference queue length is 100, and the predictive domain N_y=6.

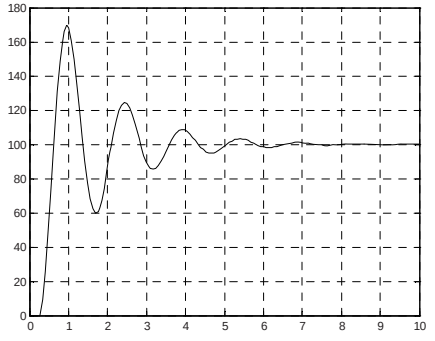


Fig. 1. The output queue length at $\lambda = 2.5$

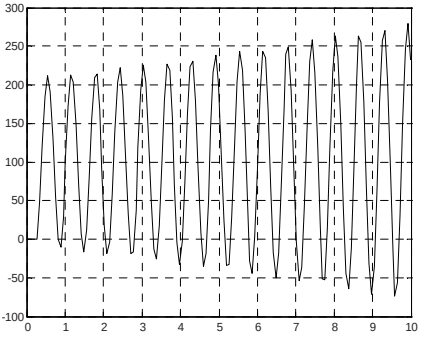


Fig. 2. The output queue length at $\lambda = 0.38$

It calculates the control parameters from above equations.

$$A(z^{-1}) = 1 - 1.7294 z^{-1} + 0.7374 z^{-2}, \quad B(z^{-1}) = (0.1255 - 0.1078 z^{-1}) z^{-4}$$

$$\sum_{j=d}^{N_y - 1} g_j (1 - a_2 g_{j+1}) = 0.17 > 0, \quad b_d + b_{d+1} = 0.0177 > 0 \quad \frac{(a_1 - 2) \sum_{j=d}^{N_y - 1} g_j^2 + \frac{1}{a_2} \sum_{j=d}^{N_y - 1} g_j}{2(1 - a_1 + a_2)} = 0.3888,$$

$$\bar{\lambda} = -0.2143 < 0$$

So when $\lambda > 0.3888$, both *H.a* and *H.b* are guaranteed, system is stability. The output queue length at $\lambda = 2.5$ is shown in Fig 1. When $\lambda = 0.38$, the output queue length is shown in Fig 2. It shows that output queue length can reach the stable level in 6 seconds in Fig 1. When $\lambda < 0.3888$, the system is instability in Fig 2.

6 Conclusion

This paper gives the discrete model and deduces expression of GPC strategy as AQM algorithm. In single-step control and predictive domain is not smaller than d+1 circumstances, it gives the closed-loop system characteristic equation. This paper uses jury criterion to prove that the system needs to meet some conditions for stability. It gives three theorems to guarantee system stability. Simulation results demonstrate the validity of our theorems.

References

1. Floyd, S., Jacobson, V.: Random Early Detection Gateways for Congestion Avoidance. IEEE/ACM Trans.on Networking 1, 397–413 (1993)
2. Misra, V., Towsley, D., Gong, W.: Fluid-based Analysis of A Network of AQM Routers Supporting TCP Flows with an Application to RED. In: Proc of the ACM / SIGCOMM 2000 Stockholm, pp. 151–160. ACM Press, Sweden (2000)
3. Hollot, C., Misra, V., Towsley, D., et al.: Analysis and Design of Controllers for AQM Routers Supporting TCP Flows. IEEE Trans Automatic Control 47, 945–959 (2002)
4. Fan, Y.F., Ren, F.Y., Lin, C.: Design A PID Controller for Active Queue Management. In: Proc.of the 8th IEEE Int'l Symp.ISCC 2003, pp. 985–990 (2003)
5. Clarke, D.W., Mohtadi, C., Tuffs, P.S.: Generallized Predictive Control. Automatica 23, 137–164 (1987)
6. Chen, Z.Q., Mao, Z.X., Du, S.Z., et al.: Analysis of PID-GPC Based on IMC Structure. Chinese J.Chem.Eng. 11, 55–61 (2003)
7. Zhang, X.H., Chen, Z.Q., Yuan, Z.Z.: Analysis and Improvement of GPC Robustness Based on Internal Model. Control and Decision 19, 542–545 (2004)
8. Kong, L.H., Chen, Z.Q., et al.: FGPC Algorithm Used in Network Congestion Control. In: Proc. of the 2005 Chinese Conference on Complex Networks, Wuhan, pp. 41–42 (2005)
9. Astrom, D.J., Wittenmark, B.: Adaptive Control, 2nd edn. Addsion Wesley Publishing Company, Boston (1994)
10. Li, T., Zhang, J.F., Chen, Z.Q.: Closed-loop Stability Analysis of Time-delayed First-order Systems under Generalized Predictive Control. Acta aeronautica et astronautica sinica 5, 678–684 (2007)

Real-Time Communications on IEC 61850 Process Bus Based Distributed Sampled Measured Values Applications in Merging Unit

Hong Hee Lee[1], Gwan Su Kim[1], Jung Hoon Lee[1], and Byung Jin Kim[2]

[1] School of Electrical Engineering,
University of Ulsan, Nam-Gu, 680-749 Ulsan, South Korea
{hhlee, gskim94, kazhan1}@mail.ulsan.ac.kr
[2] Hyundai Heavy Industries, Dong-Gu, 682792 Ulsan, South Korea
itisme@hhi.co.kr

Abstract. In recent years, IEC 61850 supports the substation automation and the standardized communication technique in both station bus and process bus. In this paper, we propose the required techniques to develop the Merging Unit, which is one of the important data acquisition equipments in substation automation, under the IEC 61850 communication protocol. Especially, we also propose the precision time-synchronization technique using IRIG-B protocol for the Merging Unit. IEC 61850-9-2 SV (Sampled Value) service is applied to the Merging Unit which is designed using embedded processor and real-time Linux operating systems. The playback program, which can reconstruct the transmitted sampled value data on the basis of time information, is developed and the performance of the IEC 61850 SV service for Merging Unit is verified experimentally.

Keywords: IEC 61850, SV, Merging Unit, Time Synchronization, IRIG-B.

1 Introduction

During the past years, the IEC TC57 has defined the IEC 61850 standard for "Communication Networks and Systems in Substations". This standard is now ready to be used not only between the station level computer and the bay level devices but also for the open communication to the primary equipment. Substations are regarded as the nodes in electric power systems. Data access and information retrieval for the power system management are come from these nodes. All protection functions are also executed similarly in the power systems. In older substations, these functions are performed by RTUs complemented by protection devices. In these days, many substations are equipped with Substation Automation Systems (SAS), which comprise all functions to be performed in substation by definition. Up to now, most of these systems are collections of devices with the dedicated functions connected only to a station level HMI and NCC gateway. Very often, it is very hard to have interoperability if the systems are composed of devices from different suppliers because of the absence of communication standard. Therefore, substation automation cannot exploit all benefits of the state-of-the-art information technology. Both their position in the power systems management hierarchy and their increased computation capacities would like to have

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1250–1257, 2008.
© Springer-Verlag Berlin Heidelberg 2008

such systems as powerful distributed preprocessors for network control systems verifying data and commands and providing the support by local automation. They may also reduce the risk for blackouts reasonably. But advanced functions complementing the existing ones and providing much more benefits have been discussed much more than the implemented. The new standard IEC 61850 elaborated by working groups of the technical committee 57 of IEC (IEC TC57) is providing the standardized communication in substation using both state-of-the-art communication technology and powerful object modeling with high-level engineering support.

Recently, IEC 61850 supports the standardized communication technique in both station bus and process bus, and presents substation automation model [1], [2]. In this paper, we propose the required techniques to develop the MU (Merging Unit), which is one of the important data acquisition equipment in substation automation, under IEC 61850 communication protocol. Especially we also propose the precision time synchronization technique using IRIG-B for the MU. IEC 61850-9-2 SV (Sampled Value) service is applied to the MU which is designed using the embedded processor and the real-time Linux operating systems. In order to evaluate the performance of the proposed MU, the playback program, which can reconstruct the transmitted sampled value data on the basis of the time information, is developed and the performance of the IEC 61850 SV service for the MU is verified experimentally.

2 Merging Unit

The MU is an interface unit that accepts multiple analog inputs such as CT/PT and digital signals and produces multiple time-synchronized serial unidirectional multidrop digital point-to-point outputs. For the 61850 communication, the MU must have the suitable structure for IEC 60044-7,-8 and IEC 61850-9-2.

2.1 Architecture of Merging Unit

The MU transfers the data using the IEC 61850-9-2 protocol after synchronizing the signals measured from the sensors installed in the electric power system with GPS. The specification of the MU is as follows:

- Transmit samples: 1200 ~ 61440 messages/sec
- Support 6 sample rates: 20, 36, 80, 256, 512, 1024 samples
- System frequency: 50Hz/60Hz
- Number of channels: 1~12 channels
- HMI user monitoring program: NI LabVIEW

The MU is divided into sensor interface, signal processing and communication part as shown in figure 1. In sensor interface part, the sensor signal is converted to the digital signal using the sampled signal received from the MU. The MU generates the sampled signals which are obtained by the predefined sampling frequency and compensates the error of time delay using the digitalized sensor data. The compensated data are transmitted to IED through IEC 61850-9-2 SV service. IED can make time equivalent data group using the date from the different MU with the aid of the time stamp obtained from GPS included in the transmitted data. The architecture of the implemented MU is shown in figure 1.

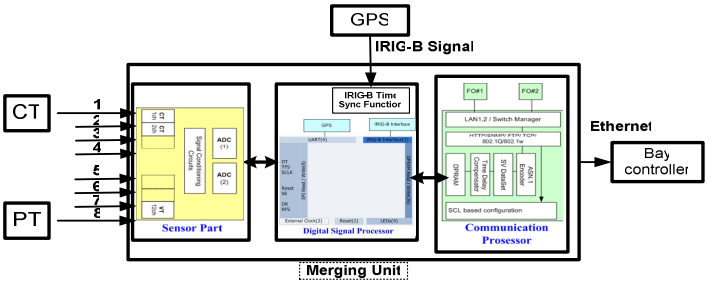


Fig. 1. Architecture of Merging Unit

Sensor Interface Part: This Part converts the measured analog signal to the digital according to the sampling command signal from the MU. The converted signal is synchronized with 1 PPS signal from the external GPS. The analog inputs from Current Transformer and Power Transformer are digitized by the A/D converter.

Signal Processing Part: The digital data through sensor interface are stored in dataset group. IEC 61850-9-2 SV service uses Publisher/Subscriber message transmission method. It composes Ethernet frame for SV message using the sampled data which are classified by dataset. When the time information is input from GPS receiver, the IRIG-B format is changed to the UTC format at every one second and the time information is stored in the UTC Time register after one second is added at the next cycle "P0".

Communication Part: The MU uses Ethernet protocol for communication. IEC 61850 communication service supports 3 kinds of services such as MMS, GOOSE and SV. The MU uses SV among these 3 services. Because TCP/IP protocol is not used in SV service under Ethernet communication, node address is assigned using MAC address in data link layer.

2.2 IRIG-B Time Synchronization

The MU transmits the measured sampled data to upper controller. So, the synchronization is essential with the upper controller. The time accuracy described in the IEC 61850-5 specification is divided into 5 classes: P1 class(1 [ms]), P2 class(0.1 [ms]), P3 class(25 [us]), P4 class(4 [us]) and P5 class(1 [us]) [3]. The time precision of a microsecond is needed for the time-synchronization of the analog data measured from the MU of the process level. The time-synchronization signal which is transmitted to the A/D converter is identical with the sampling rate. The time-synchronization of the MU is accomplished as following method; First, it happens exactly in rising edge of PPS signal. At this time, the MU transmits the converted synchronization signal to the A/D converter. If the first signal has expired or is temporarily disturbed, the MU transmits the second signal using internal clock to the A/D converter. And, if the synchronized input signal is resumed, MU carries out the time-synchronization and follows the first signal.

In 1956, the Telecommunications Working Group (TCWG) of the American Inter Range Instrumentation Group (IRIG) was mandated to create a standard format for

the distribution of synchronized time signals, and proposed the standardized set of the time code format which is described in the IRIG document 104-60. The standard was revised many times over the several years and the latest revised edition is 200-04[4]. IRIG is the industrial standard on the fact about the time synchronization of the electrical power system apparatus. The most general two formats are the modulation format (IRIG-B1xx) and the demodulation format (IRIG-B0xx). The IRIG-B modulated signal uses a 1 [kHz] carrier signal and provides a typical accuracy of 1 [ms]. And the demodulated signal can provide the precision of the nanosecond range in the case that the precision of the GPS clock device is supported. The reason why the precision of the modulated signal is smaller than the demodulated signal is due to the zero-crossing method for sensing high- and low- bit state transition [5].

The SEL-2407 GPS clock provides 100 [ns] time precision. This is used as the IRIG-B signal generator. The IRIG-B002 format which is defined in the IRIG standard is applied. The IRIG-B time-synchronization device for the MU is shown in figure 2. This device receives the IRIG-B signal from the GPS satellite clock and extracts the time information. The GPS clock transmits the IRIG-B signal to all IED nodes. By using the IRIG-B protocol which is implemented on a node, each IED nodes extracts the time information required for the time-synchronization. The IRIG-B protocol is implemented by using the Verilog-HDL language. The ALTERA FLEX 10K FPGA chip is used in order to implement the proposed algorithm. Using the Quartus II tool, the verification and the simulation of the proposed algorithm were performed.

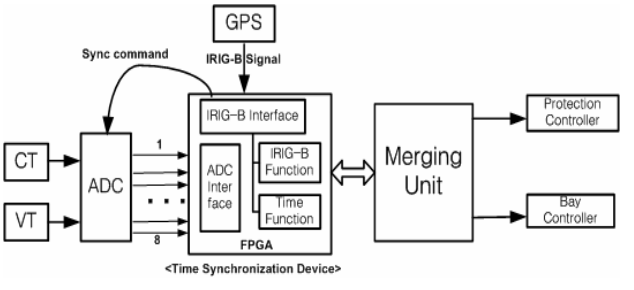


Fig. 2. IRIG-B Time Synchronization Device

The acquisition method of the time information is proposed in order to design the IRIG-B time-synchronization device using FPGA. Figure 3 shows the IRIG-B signal analysis method. In figure 3, the first, the second and the third shaded section mean the bit 0, s the bit 1 and the position value respectively. The IRIG-B signal is counted during one cycle (10[ms]) by using the internal clock. If the counter value is 3, it is interpreted as bit 0 in binary and if counter value is 6, it recognizes as bit 1 in binary. And in the case of 9, it is regarded as the position value. The time information of the BCD code form can be obtained after the 100 IRIG-B pulses are counted using the internal clock.

The basic time-synchronization experiment set is composed as shown in figure 4 in order to verify the performance of the proposed IRIG-B time-synchronization device. The message transmission with IRIG-B time-synchronization device is carried out

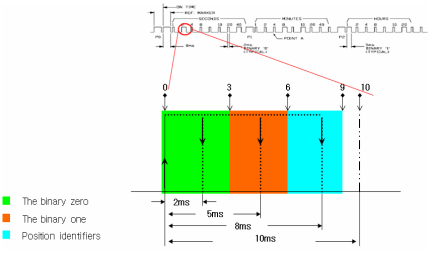


Fig. 3. IRIG-B Signal Analysis

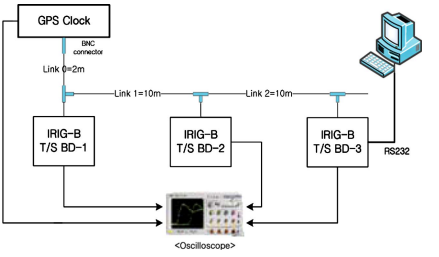


Fig. 4. Time Synchronization System

experimentally. In an experiment, the distance between the GPS clock device and the first node is set with 2 [m] and the distance of the second node and the third node is set with 10 [m].

And also, IRIG-B time code format is set with IRIG-B002 and then the experiment is carried out. The UTC time information is obtained through the IRIG-B signal received from the GPS clock device in each node and then the time-synchronization is performed. Consequently, the time precision of 200 [ns] can be identified as shown in figure 5. From the experimental result, the delay time is increased in proportion to a distance, and its magnitude is justified about 100 [ns] every 10 [m].

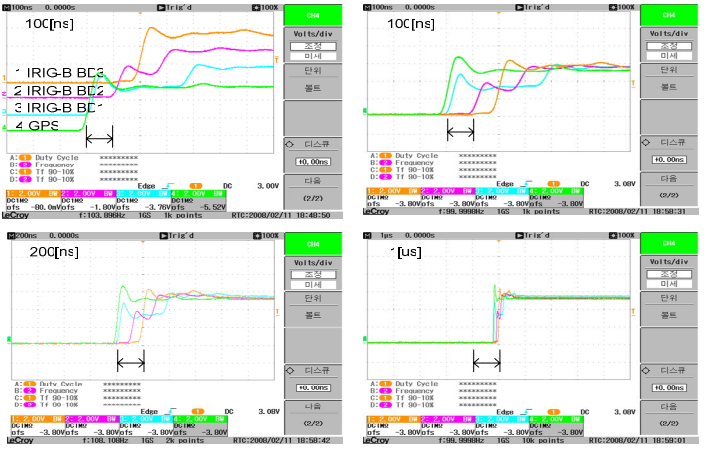


Fig. 5. Time Synchronization Delay Measured by an Oscilloscope

3 IEC 61850-9-2 SV Service

IEC 61850 SV service is one of the main communication services within a substation automation system [6]. As discussed before, the different services have different real time requirements. IEC 61850 therefore specifies that MMS communication is based on the standard TCP/IP due to its time unrestraint, while SV and GOOSE communication are directly mapped to the Ethernet layer because of the time restriction. The

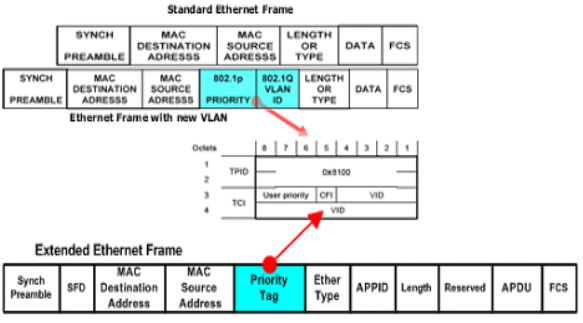


Fig. 6. SV Ethernet Frame

IEC 61850-9-2 SV service is accomplished in the designed MU. IEC 61850-9-2 SV service applies to the communication between the MU for the current transformer or power transformer and bay devices such as protection relays [7], [8].

IEC61850-9-2 SV service transmits the message encoded by ASN.1 through the multicasting to several nodes. The SV service has extended Ethernet frame structures as shown in figure 6 which illustrates extended Ethernet frame for SV message transmission.

4 Experiment and Discussion

To evaluate the performance of the proposed MU system, the basic IED communication test bed is constructed like figure 7.

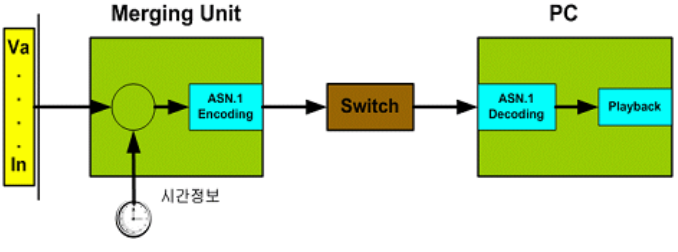


Fig. 7. Configuration of SV Communication System

In this experiment, the sampled analog data is mapped to the APDU field in the Ethernet frame. This data is received through SV message in PC and then is decoded using ASN.1. Therefore, the transmitted data can be monitored using playback program as shown in figure 8. In this playback program, LabVIEW tool of NI (National Instruments) is used. This program displays the CT and PT values which are reconstructed on the basis of the time information.

In this paper, the on-line monitoring system is also implemented, and the implemented services are as follows:

Fig. 8. Real-Time Monitoring and Measurement Item in Playback Program

- Remote monitoring and measurement
- Measurement item: voltage, current, frequency, power, THD, vector plotting
- Real-time monitoring of voltage event: Swell, Sag, Transient, Interruption

The transmission of SV message is analyzed using MMS Ethereal network analyzer. Figure 9 illustrates the process of data exchange between the PC and the MU in MMS-Ethereal.

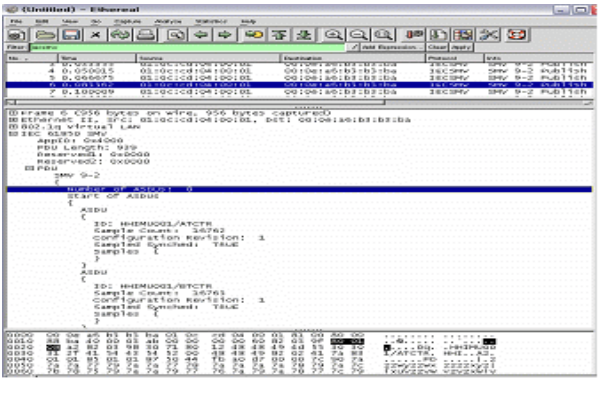


Fig. 9. SV Message Analysis using MMS-Ethereal

As can be seen from the IEC 61850 communication test between the MU and the PC, IEC 61850-9-2 SV message transmission is successfully executed experimentally. And the dataset based 8 channel /36 sampling/60 [Hz] signal is composed and it is transmitted to the upper controller using SV message frame with 100 [Mbps]. The transmitted SV message from the MU is decoded in PC and then it is reconstructed to the sampled analog data using playback program in order to verify the successful transmission of the sampled message. And the operating status of IEC 61850-9-2 SV communication service is confirmed using network analyzer.

5 Conclusions

IEC 61850 communication service defines MMS, GOOSE and SV service. Especially IEC 61850-9-2 SV service of the process level is the most basic I/O level communication service in the MU.

In this paper, the MU system was designed using embedded processor and real-time Linux operating systems. And the IEC 61850-9-2 SV communication service was achieved. In order to evaluate the performance of the proposed MU, the playback program which can reconstruct the received analog data on the basis of the time information was developed. The results of the feasibility experiments show that the proposed MU system has the enough performance for the IEC 61850-9-2 SV service.

Acknowledgment. The authors would like to thank Ministry of Knowledge Economy and Ulsan Metropolitan City which partly supported this research through the Network-based Automation Research Center (NARC) at University of Ulsan.

References

1. Apostolov, A., Auperrian, F.: IEC 61850 Process Bus based Distributed Waveform Recording. In: PES International Conference (2006)
2. Vandiver, B., Apostolov, A.: Testing of IEC 61850 Sampled Analog Values based Functions. In: Cigre 2006, Paris (2006)
3. International Standard IEC 61850-5, Communication Networks and Systems in Substations– Part 5: Communication Requirements for Functions and Device Models, First edition (2003)
4. RCC, Telecommunications and Timing Group.: IRIG SERIAL TIME CODE FORMAT. IRIG Standard 200-04
5. IEEE Standard C37.118-2005.: IEEE Standard for Synchrophasors for Power Systems
6. Engler, F., Kruimer, B., Schwarz, K.: IEC 61850 based Digital Communication as Interface to the Primary Equipment. In: Cigre 2004, Paris (2004)
7. International Standard IEC 61850-9-1, Communication Networks and Systems in Substation – Part 9-1: Specific Communication System Mapping (SCSM) – Sampled Values over Serial Unidirectional Multidrop Point-to-Point Link, First Edition 05 (2003)
8. International Standard IEC 61850-9-2, Communication Networks and Systems in Substation – Part 9-2: Specific Communication System Mapping (SCSM) – Sampled Values over ISO/IEC 8802-3, First Edition 05 (2003)

Reconstruction Algorithms with Images Inferred by Self-organizing Maps

Michiharu Maeda

Department of Computer Science and Engineering
Fukuoka Institute of Technology
3-30-1 Wajiro-higashi, Higashi-ku, Fukuoka, Japan
maeda@fit.ac.jp

Abstract. For restoring a degraded image, reconstruction algorithms with images inferred by self-organizing maps are presented in this study. Multiple images inferred by self-organizing maps are prepared in the initial stage, which creates a map containing one unit for each pixel. Utilizing pixel values as input, image inference is conducted by self-organizing maps. An updating function with threshold according to the difference between input value and inferred value is introduced, so as not to respond to noisy input sensitively. The inference of an original image proceeds appropriately since any pixel is influenced by neighboring pixels corresponding to the neighboring setting. By using the inferred images, two approaches are presented. The first approach is that a pixel value of a restored image is a median value of inferred images for respective pixels. The second approach is that a pixel value is an average value of them. Experimental results are presented in order to show that our approach is effective in quality for image restoration.

Keywords: image restoration, self-organizing map, degraded image.

1 Introduction

For self-organizing neural networks, by utilizing the mechanism of the lateral inhibition among neurons, the network with the local and topological ordering has been realized [1,2]. The neighboring neurons would always respond for neighboring inputs. For the localized inputs, obviously, the outputs react locally. Huge amounts of information are locally represented and their expressions form a configuration with topological ordering. As an application of self-organizing neural networks, there are the combinatorial optimization problem, pattern recognition, vector quantization, and clustering [3]. These are useful when there exists redundancy among input data. If there is no redundancy, it is difficult to find specific patterns or features in the data. Although a number of self-organizing models exist, they differ with respect to the field of application. For self-organizing neural networks, the ordering and the convergence of weight vectors have been mainly argued [4]. The former is a topic on the formation of topology preserving map, and outputs are constructed in proportion to input characteristics [5,6]. For instance, there is the traveling salesman problem as an application of feature maps,

D.-S. Huang et al. (Eds.): ICIC 2008, LNCS 5226, pp. 1258–1267, 2008.
© Springer-Verlag Berlin Heidelberg 2008

which is possible to obtain fine results by adopting the elastic-ring method with many weights compared to inputs [7,8]. The latter is an issue on the approximation of pattern vectors, and the model expresses enormous information of inputs to a few weights. It is especially an important problem for the convergence of weight vectors, and asymptotic distributions and quantitative properties for weight vectors have been mainly discussed when self-organizing neural networks are applied to vector quantization [9,10,11,12]. For image restoration, the smoothing methods, such as the moving average filter and the median filter, have been well known as a plain and useful approach. From the standpoint of distinct ground, the inference of original image has been conducted by the model of Markov random field formulated statistically, based on the concept that any pixel is affected by neighboring pixels [13,14].

In this study, reconstruction algorithms with images inferred by self-organizing maps are presented for restoring a degraded image. Multiple images inferred by self-organizing maps are prepared in the initial stage. By using the inferred images, two approaches are presented. The first approach is that a pixel value of a restored image is a median value of inferred images for respective pixels. The second approach is that a pixel values is an average value of them. Our model forms a map in which one element corresponds to each pixel. Image inference is conducted by self-organizing maps using pixel values as input. A renewal function with threshold is introduced in proportion to the difference between input value and inferred value. According to this function, our approach is irresponsive to input including noise oversensitively. As any pixel is influenced by neighboring pixels corresponding to neighboring setting, the inference of an original image is appropriately promoted. In the inference process, the effect of initial threshold and initial neighborhood on accuracy is examined. Experimental results are presented in order to show that our approach is effective in quality for image restoration.

2 Self-organizing Maps

Self-organizing maps realize the network with the local and topological ordering by utilizing the mechanism of the lateral inhibition among neurons. Neighboring neurons usually respond to the neighboring inputs. Huge amounts of information are locally represented and their expressions form a configuration with the topological ordering. For self-organizing maps, Kohonen's algorithm exists and is known as a popular and utility learning. In this algorithm, the updating of weights is modified to involve neighboring relations in the output array. In the vector space R^n, an input $\boldsymbol{x}$, which is generated on the probability density function $p(\boldsymbol{x})$ is defined. The input $\boldsymbol{x}$ has the components x_1 to x_n. An output unit y_i is generally arranged in an array of one- or two-dimensional maps, and is completely connected to inputs via w_{ij}.

Let $\boldsymbol{x}(t)$ be an input vector at step t and let $\boldsymbol{w}_i(0)$ be weight vectors at initial values in R^n space. For given input vector $\boldsymbol{x}(t)$, we calculate the distance between $\boldsymbol{x}(t)$ and $\boldsymbol{w}_i(t)$, and select the weight vector as winner c minimizing the

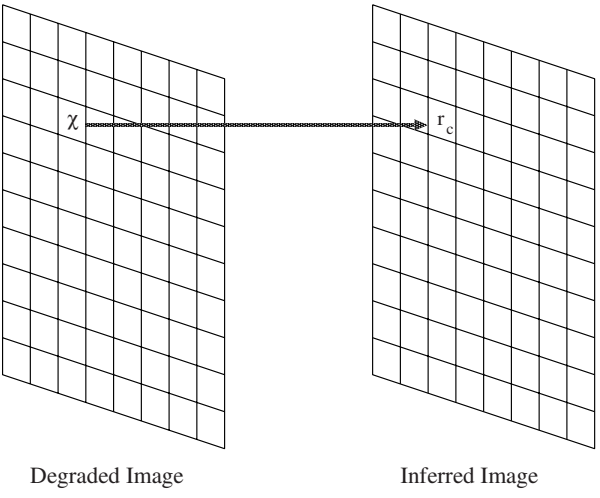


Fig. 1. Correspondence between degraded image and inferred image

distance. The process is written as follows:

$$c = \arg\min_i \{\|\boldsymbol{x} - \boldsymbol{w}_i\|\}, \tag{1}$$

where $\arg(\cdot)$ gives the index c of the winner.

With the use of the winner c, the weight vector $\boldsymbol{w}_i(t)$ is updated as follows:

$$\Delta \boldsymbol{w}_i = \begin{cases} \alpha(t)\,(\boldsymbol{x} - \boldsymbol{w}_i) & (i \in N_c(t)), \\ \boldsymbol{0} & (\text{otherwise}), \end{cases} \tag{2}$$

where $\alpha(t)$ is the learning rate and is a decreasing function of time $(0 < \alpha(t) < 1)$. $N_c(t)$ has a set of indexes of topological neighborhoods for winner c at step t.

3 Image Restoration

When self-organizing maps are adapted to the traveling salesman problem, many weights compared to inputs, about three times, are used. By disposing an array of one-dimensional map for output units, fine solutions on the basis of the position of weights after learning have been obtained approximately. In the meantime, when self-organizing maps refer to vector quantization, a few weights compared to inputs are utilized for the purpose of representing huge amounts of information, and a number of discussions have been made on asymptotic distributions and quantitative properties for weight vectors.

In this section, a learning model in relaxation algorithm influenced by self-organizing maps for image inference is presented with the same number both of inputs and weights in order to infer an original image from a degraded image

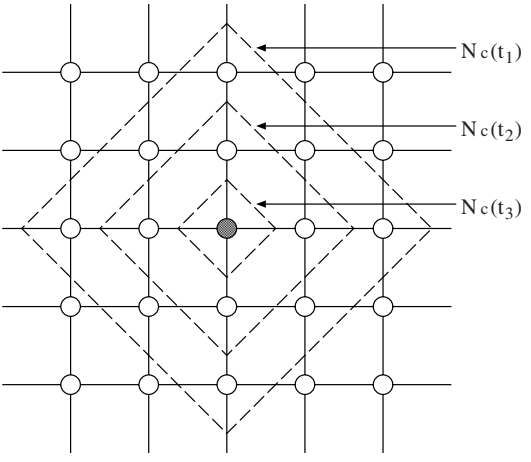


Fig. 2. Distribution of topological neighborhoods

provided beforehand [15,16]. The purpose of this study is to infer the original image from the degraded image containing random noise in sequential processing. Here, input χ as the degraded image and weight r_i as the inferred image are defined. A map forms that one element reacts for each pixel, and image inference is executed by self-organizing maps using pixel values as input.

To begin with, the value of r_i is randomly distributed near the central value of gray scale as initial value. Next, degraded image with $l \times m$ size is given. Input χ as pixel value is arbitrarily selected in the degraded image, and let r_c be a winner of the inferred image corresponding to χ. As shown in Fig.1, both of the positions χ and r_c agree under the degraded image and the inferred image. Therefore, inferred image r_i is updated as follows:

$$\Delta r_i = \begin{cases} \alpha(t)\Theta\left(\chi - r_i\right) & (i \in N_c(t)), \\ 0 & (\text{otherwise}), \end{cases} \tag{3}$$

where $\Theta(a)$ is a function in which the value changes with threshold $\theta(t)$ presented as follows:

$$\Theta(a) = \begin{cases} a\ (|a| \leq \theta(t)), \\ 0\ (\text{otherwise}). \end{cases} \tag{4}$$

$\theta(t)$ is the difference between input χ and inferred image r_i in time t and the decreasing function of time, as $\theta(t) = \theta_0 - \lfloor \theta_0 t / T_{max} \rfloor$, where T_{max} is a maximum iteration and θ_0 is an initial threshold determined by trial and error as shown in numerical experiments. In the case of learning according to self-organizing maps, weights tend to react sensitively for noisy inputs. In order to avoid the tendency, Eq.(3) is adopted, instead of Eq.(2). By applying the function, the value which obviously differs from neighboring pixels would be disregarded. Using the

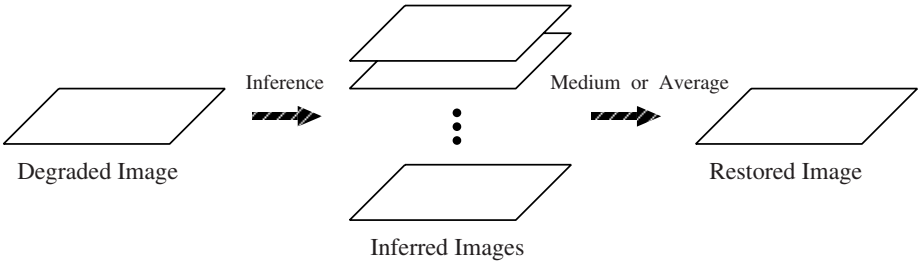


Fig. 3. Concept for image restoration. Multiple images inferred by self-organizing maps are prepared in the initial stage. By using the inferred images, a restored image is constructed by the medium model (Model I) or the average model (Model II).

functions, weights are updated until the termination condition is met. Image inference is appropriately promoted as given in the next section.

Figure 2 shows an example of the arrangement of topological neighborhoods. The circle signifies the weight and the line which connects the circles denotes the topological neighborhood. In this figure, the black circle expresses the weight of winner c. As the set of topological neighborhoods changes $N_c(t_1)$, $N_c(t_2)$, and $N_c(t_3)$ when the time varies t_1, t_2, and t_3, respectively, it is shown that the number of topological neighborhoods decreases with time. By obtaining information of the neighboring pixels, it is possible to complement lost information about pixels based on the updating function.

Image Inference in self-organizing maps (IISOM) algorithm is presented as follows.

[*IISOM algorithm*]

Step 1 Initialization:
Give initial weights $\{r_1(0), r_2(0), \cdots, r_{lm}(0)\}$ and maximum iteration T_{max}. $t \leftarrow 0$.

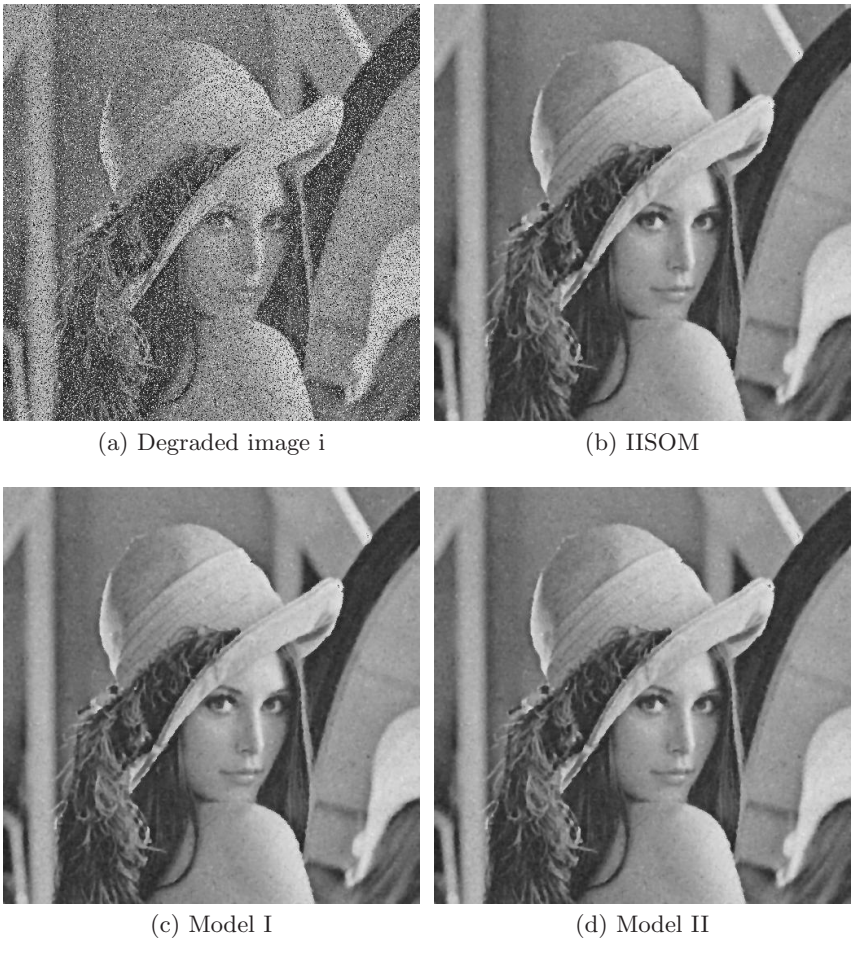
Step 2 Learning:
(2.1) Choose input χ at random among $\{\chi_1, \chi_2, \cdots, \chi_{lm}\}$.
(2.2) Select r_c corresponding to input χ.
(2.3) Update r_c and its neighborhoods according to Eq.(3).
(2.4) $t \leftarrow t + 1$.

Step 3 Condition:
If $t = T_{max}$, then terminate, otherwise go to Step 2.

In this study, a peak signal to noise ratio (PSNR) P is used as the quality measure after learning for image restoration. PSNR P is presented as follows:

$$P = 10 \log_{10}(\sigma/E) \quad [\text{dB}] \tag{5}$$

where σ and E are the square of gray-scale length, i.e., $\sigma = (Q-1)^2$ as a gray scale Q, and mean square error between the original image and the inferred image, respectively.



(a) Degraded image i (b) IISOM

(c) Model I (d) Model II

Fig. 4. Degraded image i with 512×512 size and 256 gray-scale, and results of IISOM, Model I, and Model II

By using the inferred images, two models are presented (See Fig.3). In the initial stage, from a degraded image, M images which are inferred by self-organizing maps are prepared. For respective pixels of M inferred images, let a pixel value of the restored image be (i) a median value (Model I) and be (ii) an average value (Model II). For model I, the median value has $M/2$ if M is an even number. For model II, we made a pixel value of the restored image which is rounded off to no decimal places.

4 Numerical Experiments

In the numerical experiments, image restoration is performed to infer the original image with the size 512×512 and gray scale 256. The degraded image contains

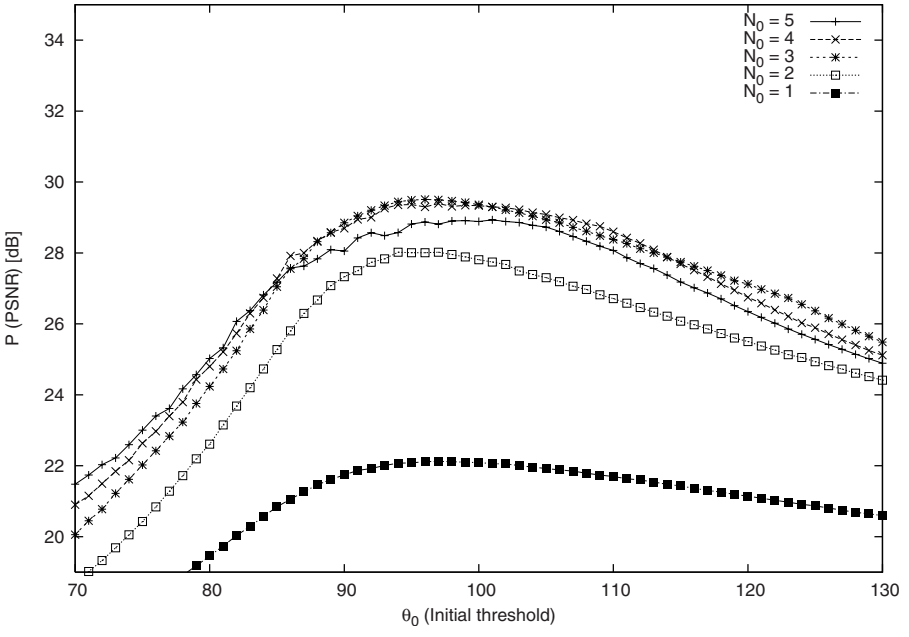


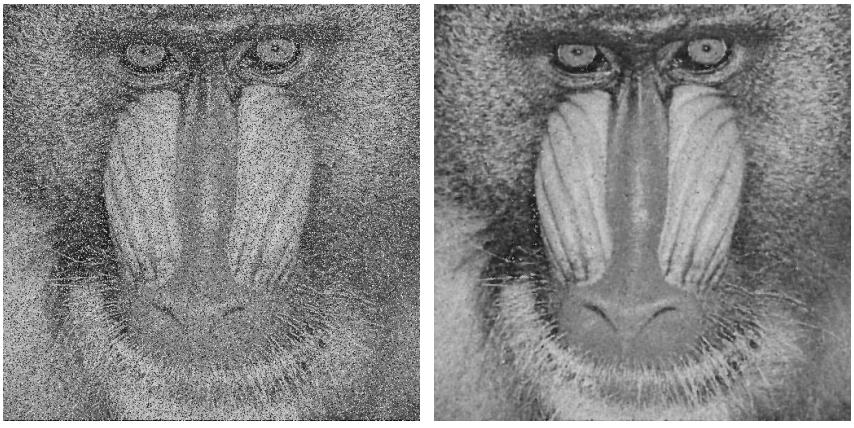
Fig. 5. PSNR and initial threshold for image i

45% noise in comparison with the original image, random-valued impulse noise, as shown in Fig. 4 (a). That is to say, the noise is included 45% pixels which are randomly chosen among the 512×512 pixels, and chosen pixels are given values from 0 to 255 at random. Initial weights are randomly distributed near the central value of gray scale Q. Parameters are chosen as follows: $l = 512$, $m = 512$, $Q = 256$, $M = 100$, $T_{max} = 100 \cdot lm$, $N(t) = N_0 - \lfloor N_0 t/T_{max} \rfloor$, and $\theta(t) = \theta_0 - \lfloor \theta_0 t/T_{max} \rfloor$.

For image restoration, Fig. 4 (b), (c), and (d) show results of IISOM, Model I, and Model II, respectively, with the initial neighborhood $N_0 = 3$ and the initial threshold $\theta_0 = 96$. According to the technique given in this study, the degraded image is restorable.

Figure 5 shows the effect of the initial threshold θ_0 on accuracy in PSNR P for each of initial neighborhood $N_0 = 1, 2, 3, 4, 5$. In this case, P yields the maximum when $N_0 = 3$ and $\theta_0 = 96$. Figure 4 (b), (c), and (d) were restored by these values.

As an example of the complicated image, Fig. 6 (a) shows the degraded image. As well as the above-mentioned image, the degraded image contains the uniform noise of 45% compared to the original image. The condition of the computation is equal to that of the earlier description. According to the present algorithm, results of IISOM, Model I, and Model II with the initial neighborhood $N_0 = 2$ and the initial threshold $\theta_0 = 103$ is shown in Fig. 6 (b), (c), and (d), respectively. It is proven that the degraded image can be also restored in this case.



(a) Degraded image ii (b) IISOM

(c) Model I (d) Model II

Fig. 6. Degraded image ii with 512×512 size and 256 gray-scale, and results of IISOM, Model I, and Model II

Table 1. PSNR for results of MAF, MF, IISOM, Model I, and Model II.(Unit: dB)

	MAF	MF	IISOM	Model I	Model II
Image i	20.4	26.2	29.5	29.7	29.8
Image ii	19.2	21.2	21.7	21.8	21.8

Figure 7 presents the effect of the initial threshold θ_0 on accuracy in PSNR P for each of initial neighborhood N_0. P yields the maximum when $N_0 = 2$ and $\theta_0 = 103$. Using these values, an example of results for image restoration is given in Fig. 6 (b), (c), and (d).

Table 1 summarizes PSNR for results of IISOM, Model I, and Model II compared to the moving average filter (MAF) and the median filter (MF). The size

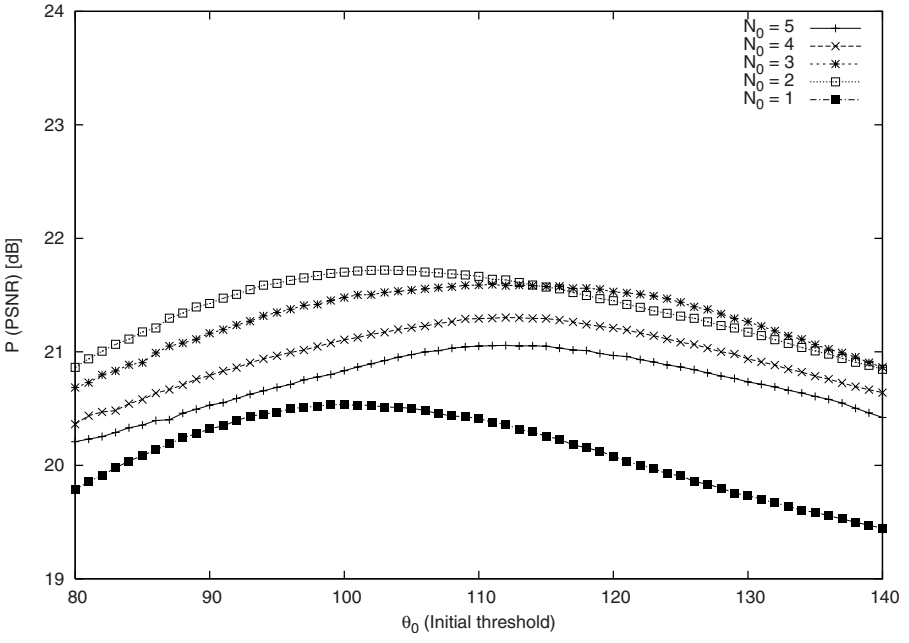


Fig. 7. PSNR and initial threshold for image ii

of the filter mask is 3×3. It is proven that Model I and Model II excel MAF, MF, and IISOM for both images i and ii.

5 Conclusions

In this study, reconstruction algorithms with images inferred by self-organizing maps for image restoration have been presented and their validity has been shown through numerical experiments. Multiple images inferred by self-organizing maps were prepared in the initial stage, which created a map containing one unit for each pixel. Utilizing pixel values as input, image inference was conducted by self-organizing maps. An updating function with a threshold according to the difference between the input value and the inferred pixel was introduced, so as not to respond to a noisy input sensitively. The inference of original image proceeded appropriately since any pixel was influenced by surrounding pixels corresponding to the neighboring setting. By using the inferred images, two approaches were presented. The first approach was that a pixel value of a restored image is a median value of inferred images for respective pixels. The second approach was that a pixel values is an average value of them. Finally, for the future works, we will study more effective techniques of our algorithms.

References

1. Grossberg, S.: Adaptive Pattern Classification and Universal Recoding: I. Parallel Development and Coding of Neural Feature Detectors. Biol. Cybern. 23, 121–134 (1976)
2. Willshaw, D.J., Malsburg, C.: How Patterned Neural Connections Can be Set Up by Self-organization. Proc. R. Soc. Lond. B. 194, 431–445 (1976)
3. Hertz, J., Krogh, A., Palmer, R.G.: Introduction to the Theory of Neural Computation. Addison-Wesley, Reading (1991)
4. Kohonen, T.: Self-organizing Maps. Springer, Berlin (1995)
5. Villmann, T., Herrmann, M., Martinetz, T.M.: Topology Preservation in Self-organizing Feature Mmaps: Exact Definition and Measurement. IEEE Trans. Neural Networks 8, 256–266 (1997)
6. Maeda, M., Miyajima, H., Shigei, N.: Parallel Learning Model and Topological Measurement for Self-organizing Maps. Journal of Advanced Computational Intelligence and Intelligent Informatics 11, 327–334 (2007)
7. Durbin, R., Willshaw, D.: An Analogue Approach to the Traveling Salesman Problem Using an Elastic Net Method. Nature 326, 689–691 (1987)
8. Angéniol, B., Vaubois, G., Texier, J.-Y.: Self-organizing Feature Maps and the Traveling Salesman Problem. Neural Networks 1, 289–293 (1988)
9. Ritter, H., Schulten, K.: On the Stationary State of Kohonen's Self-organizing Sensory Mapping. Biol. Cybern. 54, 99–106 (1986)
10. Ritter, H., Schulten, K.: Convergence Properties of Kohonen's Topology Conserving Maps, Fluctuations, Stability, and Dimension Selection. Biol. Cybern. 60, 59–71 (1988)
11. Maeda, M., Miyajima, H.: Competitive Learning Methods with Refractory and Creative Approaches. IEICE Trans. Fundamentals E82–A, 1825–1833 (1999)
12. Maeda, M., Shigei, N., Miyajima, H.: Adaptive Vector Quantization with Creation and Reduction Grounded in the Equinumber Principle. Journal of Advanced Computational Intelligence and Intelligent Informatics 9, 599–606 (2005)
13. Geman, S., Geman, D.: Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE Trans. Pattern Anal. Mach. Intel. 6, 721–741 (1984)
14. Maeda, M., Miyajima, H.: State Sharing Methods in Statistical Fluctuation for Image Restoration. IEICE Trans. Fundamentals E87-A, 2347–2354 (2004)
15. Maeda, M.: A Relaxation Algorithm Influenced by Self-organizing Maps. In: Kaynak, O., Alpaydın, E., Oja, E., Xu, L. (eds.) ICANN 2003 and ICONIP 2003. LNCS, vol. 2714, pp. 546–553. Springer, Heidelberg (2003)
16. Maeda, M., Shigei, N., Miyajima, H.: Learning Model in Relaxation Algorithm Influenced by Self-organizing Maps for Image Restoration. IEEJ Trans. Electrical & Electronic Engineering 3, 404–412 (2008)

Author Index

Ahmad, Jamil 601
Aihara, Kazuyuki 946
Akdemir, Bayram 498
Al-Rabadi, Anas N. 15, 23
Alam, M.A. 138
Aly, Saleh K.H. 1156
An, Senjian 446
Attarchi, Sepehr 356

Bai, Hua 912
Bai, Li 196
Bian, Rui 874
Bian, Yuchao 1064
Bilen, Hakan 1164
Bing, Wang 593

Cao, Hui 275
Chai, Duck Jin 1095
Chang, Shun-Chieh 227
Chang, Tun-Yu 661
Chen, C.L. Philip 696
Chen, Chien-Ta 937
Chen, Chun-Hua 661
Chen, Chuyao 688
Chen, Feng 1206
Chen, Guang-Bo 850
Chen, Guanghua 904
Chen, Hu 1064
Chen, Hui 904
Chen, Hung-Cheng 841
Chen, Jianfu 55, 430
Chen, Juan 258
Chen, Luonan 482
Chen, Mi 628
Chen, Tianding 293, 577
Chen, Wensheng 422
Chen, Yang 946
Chen, Yen-Ru 204
Chen, Yifei 970
Chen, Yuehui 79, 251
Chen, Yufeng 414
Chen, Zeng-qiang 1242
Cheng, Jian 235
Cheng, Peipei 620
Cheng, Yun-Maw 978
Cheong, Marc 1017
Chiang, Ziping 937
Chiu, Min-Sen 850
Cho, Tae Ho 1234
Chou, Chien-Cheng 937
Cichocki, Andrzej 122
Cong, Tran-Xuan 1222

Dai, Juan 696
Dashevskiy, Mikhail 380
Deng, Zhi-dong 1242
Ding, Shifei 315
Dong, Ming Chui 962
Dong, Xiangjun 372
Du, Haifeng 817
Duan, Yongrui 755

Faez, Karim 356
Fan, Dongmei 389, 397
Fan, Jian 1198
Fan, Shixuan 704
Fan, Zizhu 92
Fei, MinRui 775, 1198
Feng, Zhiquan 251
Fu, Jingqi 896

Gan, Guojun 921
Gao, Lixin 912
Gao, Mei-Juan 858
Gao, Xinjiao 167
Gao, Yongliang 179
Geng, Runian 372
Gribova, Valeriya 792
Gu, Wenjin 55
Güneş, Salih 498
Guo, Heqing 644
Guo, Ping 514
Guo, Sheng-Bo 529
Guo, Xijuan 179
Guo, Yi-nan 235

Han, Feiyu 321
Han, Kyungsook 114, 130
Hao, Yuanling 904
He, Huan 561

He, Ning 1141
He, Wei 86
He, Yuqing 561
Heo, Byeong Mun 1095
Hocaoglu, Muhammet A. 1173
HoCho, Tae 636
Hodgkinson, Ann 529
Hong, Wenxue 506
Hsieh, Chaur-Heh 337
Hu, Chi-Ping 696
Hu, Jun 529
Hu, Kai-Wen 661
Hu, Peng 696
Hu, Rong 98
Hu, Yu-Lan 521
Hu, Yulan 405
Huang, Chenn-Jung 661
Huang, Chin-Pan 337
Huang, Feng 1198
Huang, Jiacai 784
Huang, Kuo-Liang 661
Huang, Liyu 438
Huang, Tianyu 414
Huang, Ying 696
Huang, Yuangui 438
Huang, Zhihua 888
Hung, Ching-Tsung 937
Hung, Mao-Hsiung 337
Hwang, Won-Joo 585

Ibrikci, Turgay 189
Ipson, Stan S. 258

Jain, S.K. 138
Jang, Sung Ho 653
Je, Hongmo 825
Jeong, Chang-Sung 491
Jeong, Mun-Ho 1181
Jia, Li 850
Jia, Lixin 275
Jia, Yuan-Yuan 1009
Jiang, He 372
Jiang, Huilan 720
Jiang, Jianmin 258
Jiang, Peng 1206
Jiang, Tian-Zi 514
Jiang, Wen 882
Jiang, Xingdong 235
Jiang, Zhenran 86
Jiang, Zongli 1001
Jin, Changjiang 167
Jin, Fengxiang 315
Jin, Ying 712
Jing, Chai 1118
Jing, Li 243
Jo, Kang-Hyun 536, 544
Jo, Sung-Up 491
Juang, Jih-Gau 833
Jung, Kwang-Sung 1214
Jwo, Dah-Jing 227

Kang, Eunyoung 610
Kang, Hyungwoo 679
Kang, Jiquan 1103
Karabulut, Mustafa 189
Kim, ByungHee 636
Kim, Byung Jin 1250
Kim, Dae-Nyeon 544
Kim, Daijin 825
Kim, Eun-Chan 1222
Kim, Gwan Su 1250
Kim, Jisu 114
Kim, Moon Jeong 610
Kim, Taeho 536
Kim, Ungmo 610
Kim, Young-gon 985
Kong, Xiangzhen 39
Koo, Insoo 1222
Kou, Chunhai 755

Lai, Bin 1134
Lee, Eunju 610
Lee, Hae Young 1234
Lee, Hong Hee 1250
Lee, Jaetak 130
Lee, Jong Sik 653
Lee, Joong-Jae 1181
Lee, Jung Hoon 1250
Lee, Sang-goo 985
Lee, Sung-Min 1214
Lee, Taehee 985
Lee, Yang Koo 1095
Lee, Yoon-Hyung 1181
Lee, Yun-Seok 491
Leong, Ka Seng 962
Li, Bao-Sheng 1149
Li, Fengxia 414
Li, Guangting 1118
Li, Guo-Zheng 106, 482

Li, Hao 474
Li, Heng 800
Li, Jing 506, 896
Li, Jinsheng 430
Li, Lei 552, 569, 874
Li, Meili 755
Li, Qi-Peng 71
Li, Tao 364
Li, Wei 79, 1087
Li, Xiaohu 817
Li, Xueling 146
Li, Yi 251
Li, Yingshan 55
Li, Yiping 962
Li, Yourong 669
Li, Yuan 808
Li, Zhibin 86
Liang, Zhao 593
Lin, Jyh-Dong 937
Lin, Yezhi 744
Lin, Yong 235
Linner, David 763
Liu, Bin 243
Liu, Chenyong 1111
Liu, Ergen 92
Liu, Feng 970
Liu, Fenlin 243
Liu, Jin 1087
Liu, Libing 1040, 1072, 1078
Liu, Lin 1047
Liu, Qian 688
Liu, Shu-bo 1087
Liu, Wanquan 446
Liu, Weiling 1040, 1072, 1078
Liu, Xiao 704
Liu, Xiaodong 446
Liu, Yihui 196
Liu, Yue 106
Liu, Yuhua 628
Loke, Kar-Seng 1017
Lu, Chong 446
Lu, Donghui 704
Lu, Jianxia 888
Lu, Ke 1141
Lu, Qin 993
Lu, Yi-Cheng 833
Lu, Yong 669
Lu, Zhimao 389, 397
Luo, Yun-Cheng 661
Luo, Zhiyuan 380

Ma, Jinwen 552, 569
Ma, Jun 474, 728
Ma, Qi 1124
Ma, Shiwei 904
Ma, Zhi 47
Maeda, Michiharu 1258
Maeda, Sakashi 1156
Manderick, Bernard 970
Mandic, Danilo P. 122
Mei, Liquan 620
Moon, Cheol-Hong 1214
Mu, Zhichun 460

Naeem, Muhammad 601
Niu, Hai-Jing 514
Nosrati, Masoud S. 356

Oh, Eun-Tack 1214
Oran, Bülent 498

Pai, Chen-Yu 978
Pan, Quan 71
Pan, Zihao 55
Pao, Tsang-Long 978
Park, Jong-Heung 985
Park, Sun 1025
Peng, Dian-Jun 31
Peng, Hui 808
Peng, Sheng-Lung 204, 212
Pfeffer, Heiko 763
Przybyszewski, Andrzej W. 122

Qi, Xuerui 704
Qin, Li-Juan 521
Qin, Lijuan 405
Qin, Renchao 364
Qingxian, Pu 585
Qiu, Daowen 1
Qu, Wei 1009

Rehman, Habib ur 601
Ren, Chun-Xiao 474
Ren, Jinchang 258
Ren, Yuan 1103
Rutkowski, Tomasz M. 122, 300
Ryu, Keun Ho 1095

Seeja, K.R. 138
Shah, Syed Ismail 601
Shan, Dapeng 720
Shang, Li 220

Shao, LiKang 1198
Shen, Lincheng 808
Sheng, Danghong 784
Shrestha, Rojan 114
Shu, Hua-Zhong 1149
Shuai, Dianxun 266
Si, Gangquan 275
Song, Yonghua 1032
Steglich, Stephan 763
Su, Jindian 644
Sun, Jia 364
Sun, Ke 1056
Sun, Mingxuan 736
Sun, Xiuli 704

Takaishi, Tetsuya 929
Tan, Xi 347
Tang, Chuan Yi 212
Tang, Jianliang 422
Tao, Jianhua 482
Tao, Peng-Ye 850
Teng, Zaixia 106
Tian, Jing-Wen 858
Trinh, Hoang-Hon 544
Tsay, Yu-Wei 204, 212
Tsuruta, Naoyuki 1156
Tu, Dawei 888

Unel, Mustafa 1164, 1173

Wan, Hong 1118
Wang, Bin 9, 1149
Wang, Bo 321
Wang, Haiying 63, 159
Wang, Hengyu 552
Wang, Hong 405, 521, 954
Wang, Hongyuan 98
Wang, Jie 422, 954
Wang, Jinjia 506
Wang, Kuisheng 1064
Wang, Lin 808
Wang, Ling 896, 1095
Wang, Qing 329
Wang, Shulin 146
Wang, Sunan 817
Wang, Tai-Chun 212
Wang, Ting 1190
Wang, Wan-Cheng 866
Wang, Wan-Liang 31
Wang, Weiming 744
Wang, Xianming 1118
Wang, Yutian 39
Wang, Zhigang 669
Wang, Zhihao 167
Wang, Zhonghua 251
Wei, Ying-Zi 521
Wei, Yingzi 405
Wen, Guangrui 712
Wen, Xiulan 784
Wong, Fai 962
Wong, Kam-Fai 993
Wu, Haitao 720
Wu, Jianguo 775
Wu, Jinhua 800
Wu, Qi 106
Wu, Tao 696
Wu, Tsung-Hsien 661
Wu, Yong 321
Wu, Yuqiang 39

Xia, Yunqing 993
Xiao, Han 669
Xie, Lei 179, 347
Xie, Zhaoxia 460
Xu, Anping 1072
Xu, Kaihua 628
Xu, Ruifeng 993
Xu, Wenbo 372
Xu, Xiao-Yan 696
Xu, Yun 167
Xu, Zheng-quan 1087
Xue, Anke 1206
Xue, Yu 167

Yang, Bo 251
Yang, Hongying 561
Yang, Qi 882
Yang, Qiang 874
Yang, Weidong 1040, 1072, 1078
Yang, Yingyun 1056
Yang, Yu-Pu 593
Yang, Zeqing 1078
Yang, Zijiang 921
Yao, Xuebiao 167
Ye, Long 1056
Yeh, Jun-Heng 978
Yi, Chucai 98
Yin, Jianan 712
Yin, Yafeng 106
Yin, Yi-Long 474

Yong, Zeng 593
Yoo, Seung-Hun 491
Yoo, Taihui 130
You, Bum-Jae 1181
You, Mingyu 482
Yu, Qilian 1040
Yu, Shanshan 644
Yuan, Zhi-yong 1087
Yuan, Zhi Yong 1134
Yue, Sicong 329

Zaka, Imran 601
Zdunek, Rafal 300
Zeng, Weimin 904
Zeng, Zhenbing 744
Zeng, Zhi-yuan 466
Zhang, An 882
Zhang, Deng Yi 1134
Zhang, Fan 858
Zhang, Haiyan 704
Zhang, Jing-Ling 31
Zhang, Ke 308
Zhang, Liming 1124
Zhang, Liqing 285
Zhang, Liyan 1064
Zhang, Nan 1032
Zhang, Peijian 775
Zhang, Peng 1141
Zhang, Qiang 9
Zhang, Qin 1056
Zhang, Qiuwen 47
Zhang, Rubo 389, 397
Zhang, Ruchuan 1190
Zhang, Rui 9
Zhang, Shanwen 146
Zhang, Shao-Wu 71
Zhang, Shouhong 438
Zhang, Wenshuai 47
Zhang, Xiaoli 904
Zhang, Xing-hui 1242
Zhang, Xingming 55, 430
Zhang, Xining 712
Zhang, Yanbin 275
Zhang, Youan 800
Zhang, Yu 364, 529
Zhang, Zhen 1103, 1111
Zhao, Hongchao 1190
Zhao, Jian Hui 1134
Zhao, Qijie 888
Zhao, Rongchun 329
Zhao, Xu 1001
Zhao, Yan-Wei 31
Zhao, Yaou 79
Zhao, Yuzhong 167
Zhen, Tong 47
Zheng, Chunhou 39
Zheng, Huiru 63, 159
Zheng, Yanwei 251
Zhi-Mei, Niu 1149
Zhou, Han 466
Zhou, Jian-zhong 466
Zhou, Shi-Ru 858
Zhou, Tie Hua 1095
Zhou, Ying 1103, 1111
Zhou, Yue 405, 521
Zhu, Daqi 688
Zhu, Tianqiao 438
Zhu, Wenjun 285
Zhu, Yuhua 47
Zhuang, Jian 817
Zhuo, Hankui 874
Zou, Kuan-sheng 1242