Mark Zhou
Honghua Tan (Eds.)

Communications in Computer and Information Science 202

Advances in Computer Science and Education Applications

International Conference, CSE 2011
Qingdao, China, July 2011
Proceedings, Part II

Part 2

Springer

Communications
in Computer and Information Science 202

Mark Zhou Honghua Tan (Eds.)

Advances in Computer Science and Education Applications

International Conference, CSE 2011
Qingdao, China, July 9-10, 2011
Proceedings, Part II

Springer

Volume Editors

Honghua Tan
Wuhan Institute of Technology
Xiongchu Road 693
Wuhan, China
E-mail: honghua_tan@yahoo.cn

Mark Zhou
Education Society
The Sherwood No.8 Fuk Hang Tsuen Road
Tuen Mun, Hong Kong
E-mail: zhoumarkhk@sina.com

ISSN 1865-0929 e-ISSN 1865-0937
ISBN 978-3-642-22455-3 e-ISBN 978-3-642-22456-0
DOI 10.1007/978-3-642-22456-0
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011931540

CR Subject Classification (1998): I.2, D.2, H.4, I.4, F.1, H.3

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The present book includes extended and revised versions of a set of selected papers from the 2011 International Conference on Computer Science and Education (CSE 2011), held during July 9–10, 2011, which was jointly organized by Springer and the International Science and Engineering Center, Hong Kong.

Qingdao, located in the southeastern part of Shandong province, is a beautiful seaside city with clean air and enchanting sea views. This city, bordered by the Yellow Sea on two sides, has the largest bathing beach in Asia and produces the mellow Qingdao Beer. To the east of Qingdao, a short distance across the Yellow Sea, lie Korea and Japan, making the city an important port for international trade. For 33 years, up to 1949, Qingdao was a colony of Germany and Japan. Thus, Qingdao has a great deal of European architecture. This foreign architecture—a remnant of the colonial past—now beautifies this seaside city. The red roofs, green trees, blue sea, and azure sky form a bright and colorful scene. The city is sometimes known as the Switzerland of the Orient.

Surrounded by the sea on three sides, Qingdao attracts many tourists with its charming seascape. Sights like the Eight Passes Villas, Wusi Square, Zhan Bridge, Mt. Laoshan, No.1 Bathing Beach, Badaguan Scenic Area make Qingdao unforgettable. Just walking on the sand and listening to the surf breaking on the beach becomes enjoyable.

The goal of CSE 2011 was to bring together researchers from computer science and education as well as practitioners to share ideas, problems and solutions relating to the multifaceted aspects of computers and education.

Being crucial for the development of computer science and education, our conference encompasses a large number of research topics and applications: from artificial intelligence to computers and information technology; from education systems to methods research and other related topics. Our expert reviewers from home and abroad ensure we have high-quality international conference through their rigorous reviewing process. All accepted papers are published the *Communications in Computer and Information Science* series by Springer.

Have been possible without the help of the Program Chairs, organization staff, and the members of the Program Committee. We thank them for their hard work.

We are confident that the conference program provided attendees with detailed insight into the new trends in this area, and we look forward to meeting you at future events.

Honghua Tan

Organization

General Chairs

Honghua Tan	Wuhan Institute of Technology, China
Junwu Zhu	Yangzhou University, China

Program Chairs

Chang Wen	Chienkuo Technology University, Taiwan
Chris Marry	Harokopio University of Athens, Greece

International Program Committee

Lizhe Wang	Virginia Polytechnic Institute and State University, USA
Benmei Chen	Gwangju Institute of Science and Technology, Korea
Dehuai Zeng	Shenzhen University, China

Publication Chairs

Honghua Tan	Wuhan Institute of Technology, China
Xiaolong Wang	Wuhan Insititue of Technology, China

Table of Contents – Part II

Table of Contents – Part I

Design and Implementation of Intelligent Light Compensation LED Light-Control System

Heng Yu, Peng Zhou, and Shi-long Zhang

School of Information Engineering, Wuhan University of Technology,
Wuhan 430070, China
125457273@qq.com

Abstract. This article uses the photo resistance collecting light intensity signal, uses the F05P and the J04V wireless transceiver modules for communication, uses the low power consumption MSP430F169 microprocessor for controlling and data processing, uses BP2808-constant current controlling chip which can convert the PWM signal outputed by MCU to the drive current controlled by the PWM duty cycle , coordinated the PID closed-loop control algorithm to design one kind of new energy-efficient light compensation light-controlling system. This system can make the best of daylight for compensation for self-luminous to make the sum of LED's brightness and the brightness of natural light keep stably in the setting brightness, so as to achieve the purpose of energy saving.

Keywords: light intensity signal; communication; PWM; PID; energy saving.

1 Introduction

In China, the consumption of electricity energy used for lighting is about 10% of the total generating capacity every year, the saving of electricity for lighting is extremely potential. Currently, Home, office and public places generally use the manual switch lighting. It's very frequent that the lights are still working when the natural light is quite strong, which results in a lot of energy waste. In addition, the unnecessary use will reduce the life of the lamp. Faced with these problems, the design uses high brightness and low power consumption LED lights as lighting devices[1], low power consumption MCU MSP430 as the processor. The design uses the natural light to compensate for self-luminous,which makes the brightness of LED adjust automatically from the brightness of the environment indoor .We also can change the lighting levels according to the needs with the remote control ,to optimize electricity energy use, which makes life more comfortable while saving energy largely.

2 Principles of the System

The system consists of two parts: smart dimmer and intelligent remote controller. Among them, the smart dimmer includes: MSP430F169 processor, dimming module, light detection module, power module and wireless receiver module; the remote control includes: MSP430F169 processor, transmitter module, key module and OLED module. Block diagram is shown in figure 1.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 1–6, 2011.

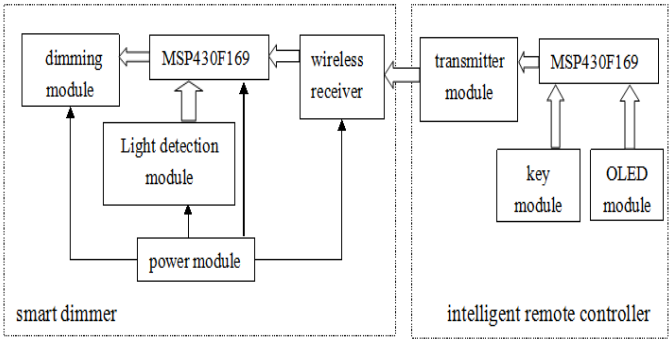


Fig. 1. Block diagram of dimming system

The system works: The wireless receiver module receives the scene data from the transmitter module, and send the data to the processor via serial port. According to the received data, the microcontroller sets a value of the corresponding scene brightness and compares the value with the data detected by light detection module to control the LED lights to adjust its brightness, making the sum of the LED lights' brightness and the brightness of natural light keep the scene brightness value set by the processor. The system compensates LED illumination using the natural light to reduce the system power loss and achieve the purpose of energy saving.

3 System Hardware Design

3.1 Hardware Design of Smart Dimmer

1) MCU control module:System uses MCU MP430F169,which integrates AD converters and serial modules in it[2].MCU compares the set value with the data detected by light detection module to chang the duty cycle of the output square wave,which results in the change of the drive current of LED lights[3], and LED lights may change the brightness in pace with it.
2) Power module: the part obtains DC power 15V (BP2808) and 3.3V (MCU) through rectification, filtration and voltage stabilizing.
3) Light detection module:this part uses the photoresistance detects luminous intensity signal, which is passed to the MCU.
4) Dimming module:This design uses a LED drived constant-current controlling chip BP2808,which reuses DIM pin for LED analog dimming, PWM dimming and dynamic temperature compensation.This design uses PWM dimming technology to control the brightness of LED lights. When changing the duty cycle of PWM inputted to control the MOSFET drain current, which is the drive current of the LED lamps, the brightness of LED lamps will change. The greater the duty cycle is, the brighter the LED lamps will be.
5) Wireless receiver module:This module uses J04V.

3.2 Hardware Design of Intelligent Remote Controller

1) MCU control module:This module uses MCU MP430F169.
2) OLED module:This module uses the low-voltage, low power consumption, small size, low cost OLED12864 LCD.
3) Transmitter module: Wireless transmitter module uses F05P. The transmission power is less than 10mW and the voltage ranges from 3 V-12V. The communication between transmitter module and the MCU is serial communication.

4 System Software Design

4.1 The Protocol Followed by the Wireless Transceiver

Provisions of this protocol: the data sent and received contains 7 bytes,which are the header-error-check("SS" is two bytes), address code(one byte), two data code (two bytes), tail checksum ("PP" is two bytes). Diagram is shown in Figure 2. Remote controllor selects different receiving ends for communication through different address code.When the header-error-check received is "SS" and the tail checksum is "PP", while data code 1 and data code 2 is the same,the receiver is thought have received the correct data. With this simple protocol the interference from the communication transceiver to other receiver-side can be reduced, and system expansion is convenient.

S	S	address code	data1	data2	P	P

Fig. 2. Structure of the data transmitted

4.2 Software Design of Smart Dimmer

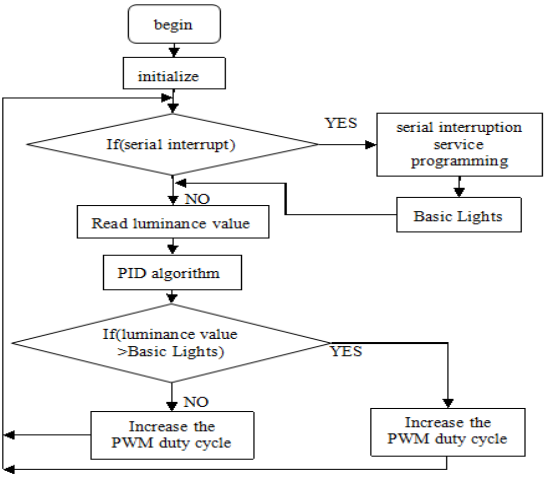


Fig. 3. Program flow diagram of smart dimmer

This part can set the initial brightness, read the ambient brightness.Comparing the ambient brightness and the initial brightness, this part can achieve automatic dimming.When the ambient light becomes strong, the brightness of LED lamps will decrease.When the ambient light becomes weak, the brightness of LED lamps will increase. You can set different scene modes so that the sum of ambient brightness and the brightness of LED lamps can be set manually, so that the environment keeps the ideal lighting condition. Program flow diagram is shown in Figure 3.

Due to ambient brightness is not constant, is in the constantly changing state, so we use the PID loop control algorithm[4][5], with the hardware circuit, making the brightness of the system more stable .The control law of discrete PID controller is:

$$\mu_k = K_P e_k + K_I \sum_{j=0}^{k} e_j + K_D (e_k - e_{k-1}) + \mu_0 \tag{1}$$

This article adopts Incremental PID control algorithm, that is to calculate $\Delta\mu$, and

$$\Delta\mu = \mu_k - \mu_{k-1} \tag{2}$$

By (1) have

$$\Delta\mu_k = Ae_k + Be_{k-1} + Ce_{k-2} \tag{3}$$

and

$$A = K_P (1 + \frac{T}{T_I} + \frac{T_D}{T}) \quad , \quad B = -K_P (1 + \frac{2T_D}{T}) \quad , \quad C = K_P \frac{T_D}{T} \tag{4}$$

Program flow diagram of PID control algorithm is shown in Figure 4.

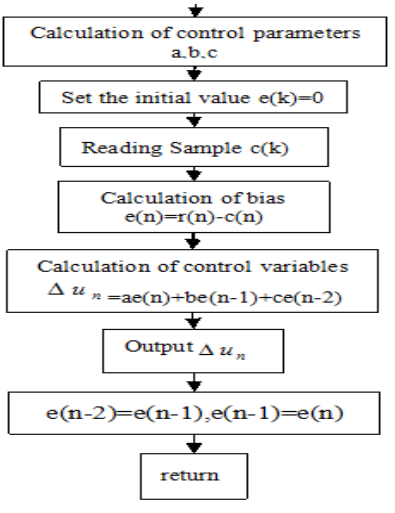


Fig. 4. Program flow diagram of PID control algorithm

4.3 Software Design of Intelligent Remote Controller

Program flow diagram of intelligent remote controller is shown in Figure 5.

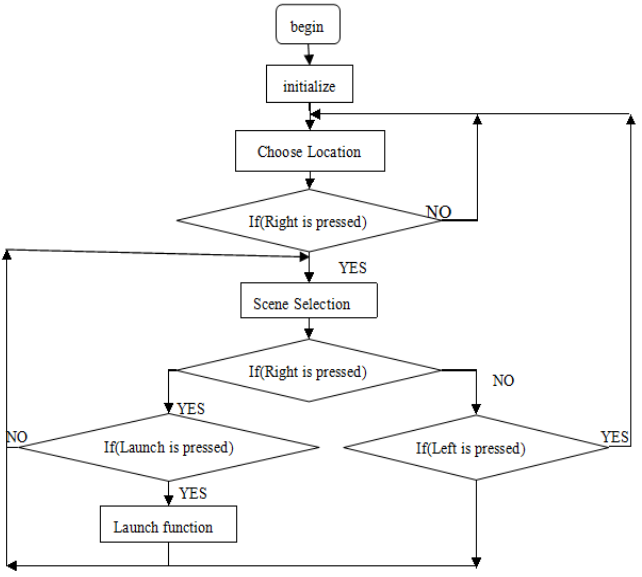


Fig. 5. Program flow diagram of intelligent remote controller

5 System Power Analysis

Table 1. Power analysis of smart dimmer[6]

Energy consumption components	MCU	J04V	BP2808	Total power consumption	Units
Energy consumption	0.45	0.52	60	61	mW

Table 2. Power analysis of intelligent remote controller

Energy consumption components	MCU	F05P	OLED12864	Total power consumption	Units
Energy consumption	0.08	0.5	67.5	68.08	mW

Note: Since the proportion of the time of standby and working is different in the two parts, the power consumption of MCU in the two parts is different.

As it can be seen from the data above, not only does the system meet the needs of the design function, but also the power consumption is quite low. Combined with the low power consumption and high brightness LED lighting,there is a significant reduction in electricity for lighting. In today's increasingly tight energy crisis, this system is highly competitive.

6 Conclusion

This system presents a practical new energy-saving design. On one hand it uses low-power chips and low-power lighting devices, on the other hand it uses clever hardware design and PID loop control algorithm, so that the system can closed-loop compensate the LED lighting according to ambient brightness to make the environment remain at a constant brightness, and can set different brightness freely according to the needs. It can save energy from many aspects and make more comfortable life. It makes up for the shortcomings of previous lighting system. Above all the system has important economic value and market prospects.

References

1. Zhou, Z.: LED lighting technology and engineering. China Electric Power Press, Beijing (2010)
2. Texas Instrument. MSP430x1xx Architecture Summary (2007)
3. Shen, J.: Low-power 16-bit microcontroller MSP430 Family Principles and Applications. Tsinghua University Press, Beijing (2001)
4. Zhu, Y., Cui, R.: Computer Control Technology, 2nd edn. Electronic Industry Press, Beijing (2010)
5. Dong, J., Zhao, C., et al.: Control Engineering, 3rd edn. Tsinghua University Press, Beijing (2010)
6. Texas Instrument. MSP430x1xx Family User’s Guide (2007)

Explorations and Practices on Nanchang University "3+1" Software Talents Cultivation Mode

Jun Hu[1] and Chun Guan[2]

[1] School of Software, Nanchang University, Nanchang, Jiangxi, China
[2] School of Information Engineering, Nanchang University, Nanchang, Jiangxi, China
hujun@ncu.edu.cn

Abstract. To cultivate the practical, Inter-disciplinary and international software talents, The School of Software, Nanchang University (SSNCU), based on the long-term practices, reformed the traditional cultivation mode of software talents, explored positively and set up "3+1" Software Engineering Professional Talents Cultivation Mode. During the process of cultivation, it divides the academic process into two stages, which includes 3-year principle theoretic teaching and one-year enterprise engineering practice teaching. To make engineering practice teaching outstanding, to focus on SSNCU-enterprises cooperation and international exchanges, SSNCU has set up more than 20 domestic and overseas practice training bases, and has gradually formed some Characteristic teaching processes such as Suite Teaching, Bilingual Teaching, Workroom Practice, Domestic SSNCU-Enterprises Training and etc. Therefore, it improves students' engineering practice abilities and achieves good results.

Keywords: Software Talents, Cultivation Mode, Enterprise Practices.

1 Introduction

With the rapid development of software Industry, the demand of software talents is also dramatically increasing. That is to change the situation of the deficiency of the talents in software industry, our country set up 35 national demonstrative software colleges in 2005, and encouraged every province, city and autonomous region with the relative policies to set up the provincial demonstrative software colleges. Several years later, with the graduation of students from various software colleges, the quantity of domestic software talents increased very rapidly. However, it was noticed very soon that the students who cultivated recently, esp. those who graduated from provincial institutes couldn't adapt to various research work of software industry, but only could engage some low-level software sale, training and maintenance work. The high-level research software talents were still in a situation of deficiency, which still seriously hindered the development of our national IT enterprises. To investigate its reasons, the main reason is that the teaching mode of the relative majors of our national computer software, even the teaching of some software institutes still followed basically the traditional cultivation mode for science & engineering talents, and some software institutes only still copied the old computer department in the college, lacked of the training for students' engineering research abilities and

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 7–13, 2011.

professional qualities, and lacked of the teaching environment to cultivate students. Therefore, the students cultivated by this mode couldn't meet the requirement of software industry, and couldn't adapt to the running of modern IT enterprises very easily.

School of Software, Nanchang University (SSNCU), which was set up in June 2002, is the demonstrative software college in Jiangxi Province. The running purpose is to cultivate international practical and Inter-disciplinary senior software talents who have good political qualities and moral cultivation, solid computer basic knowledge, regular software design and development abilities, good organization and communication abilities with international comprehensive qualities, and can adapt to domestic environment. With many years continuous exploration of all faculty in SSNCU, through revolution and innovation of brand-new running mode, curriculum system, teaching content and teaching methods, focusing on the engineering practice evaluation, strengthening the practices of principle software establishment qualities, and the cultivation of abilities in logical thinking, scientific analysis method, emphasizing on the cultivation of students' systemic design and analysis and software engineering program management as the running characteristics, implementing Theoretical Teaching, Technology Training, Engineering Practices, Bilingual Teaching, Suite Teaching and etc., SSNCU has cultivated a lot of practical talents who have not only technology development abilities but also engineering management abilities.

During the process of formulating 2008 Nanchang University Cultivation Plan, all SSNCU faculty went all out to summarize the teaching characteristics, experiences and lessons in the past 6 years, and the experiences and lessons. Besides, SSNCU developed teaching plan, decreased theoretical curriculums, strengthened engineering practice and promoted the master of theoretical knowledge during the engineering practice process. In the final Cultivation Plan, the theoretical curriculums should be completed within the first 3 years, and the first term of the last year should be spent for enterprise engineering practice, and the second term of the last year should be spent for the graduate internship and graduate design, that is, to adopt "3+1" teaching mode. What deserves us to note is, The Outstanding Engineers Cultivation Plan put forward by our country in 2009 is just the cultivation mode of "3+1", which is the same cultivation mode in our 2008 Cultivation Plan by coincidence, therefore, it proved the Sophistication and accuracy of our new edition of SSNCU Cultivation Plan.

2 New Cultivation Mode of "3+1" Software Talents

The new Cultivation mode of "3+1" Software Talents, which executed in the 2008 edition of SSNCU Teaching Plan, has been formed through many years' practices and explorations. Since the foundation of SSNCU in 2002, SSNCU has emphasized on the academic characteristic of engineering practice training, at first, SSNCU assigned teaching hours every week for engineering practice teaching in every semester; then during the process of execution, we found that we should adopt different engineering practice methods for different engineering practice courses during different stages or periods. Some can adopt the old execution methods, some should be learnt for one to

two weeks by the end of the semester, and some should be executed one semester earlier before doing the graduation design and internship. The software engineering development abilities of SSNCU students obviously are much better than those who graduated from the computer department of the same level universities or software colleges. According to this arrangement, SSNCU has formed a new cultivation mode of "3+1" Software Talents.

"3+1" mode was put forward formally during the process of editing the 2008 Nanchang University Teaching Plan. The undergraduates who enrolled their names after 2008 had to fulfill overall this teaching plan completely. What we mean "3+1" is to divide the undergraduate teaching into two stages, one is the first 3 years, the other is the last one year. These two stages could be called engineering practice teaching stage in the school and enterprises engineering practice teaching stage. Whatever the teaching in the school or outside of school, SSNCU emphasizes on the characteristic of engineering teaching, and regards engineering quality cultivation as the main line from the beginning to the end during the process of cultivating students. For the first 3 years' engineering practice teaching stage in the school, on one hand, from the third semester to the sixth semester, in every semester, SSNCU sets one engineering practice training course which mainly focuses on cultivating students' engineering abilities and professional qualities; on the other hand, SSNCU makes a engineering reformation for the contents of more than 10 traditional courses, such as Data Architecture and Algorithms, Operation System, Database System, Multi-media Technology and etc. Therefore, SSNCU enriches the content of engineering teaching and emphasizes the practical content in enterprises. For the last year' s enterprises engineering practice teaching stage, students have four choices, that is, Firstly, join the doctoral workroom in the school, to work and study with the teachers in the workroom and to take part in the project and program research ; Secondly, join the engineering practice teaching class which set up in the school by SSNCU and enterprises, to learn how to develop the program on site from the enterprise instructor; thirdly, to work in a internship organization contacted by SSNCU and participate in the project development; fourthly, Students contact the internship organization himself or herself. Whatever choice the student take, the content of program development which the student participated in should be in accordance with the training content specified by the teaching plan, or the student couldn't get the scores for engineering practice training course. Absolutely, the content range of engineering practice specified by the teaching plan is very extensive, generally speaking, only if the student takes part in one of the four program development methods mentioned above, the student could gain better engineering practice training.

3 Characteristics of "3+1" Software Talents Cultivation Mode

During the process of execution the new mode of "3+1" Software Talents, SSNCU insisted on the characteristics of Engineering Practice Teaching, Suite Teaching, Bilingual Teaching, Workroom Practice, Domestic and Overseas Enterprises Training. All these characteristic teaching ensured the successful execution of new mode. There are some brief introductions to these teaching characteristics as follows.

3.1 Engineering Practice Training

In the new edition of new cultivation mode of "3+1" Software Talents, SSNCU organically adds Engineering Practice Training Teaching courses, mainly distributes from the third semester to the seventh semester, among these five semesters, from the third semester to the sixth semester, one course for every semester, but in the seventh semester, SSNCU executes several engineering training practice courses with different methods according to different major direction. When SSNCU arranges the teaching plan every semester, every department and every major should submits the teaching schedule for that semester, and the content of setting every semester should make some adjustment according to the teaching plan, engineering practice plan and the teaching schedule for that semester. The evaluation of engineering practice training normally adopts checking method, and every course has different evaluation method. Generally speaking, Engineering Practice Training doesn't adopt the checking method, such as paper examination, but requires students finish some big homework or master the relative operating skills.

Engineering Practice Training Teaching of SSNCU has made certain progress, and has won a good fame among students and all walks of life for these years.

3.2 Suite Teaching

Suite Teaching, also called comprehensive teaching, is designed, to meet the requirement of enterprises. SSNCU integrates a lot of relative courses together, including both theory knowledge and engineering practice, especially focuses on evaluating the study effectiveness of students with comprehensive course design.

Until now, the Suite Teaching courses we have set include: 1. Computer Architecture, which includes sixth courses, that is, Computer Principle, Computer Architecture, Digital Logic, Digital Circuit, Principles of Single-chip Microcomputer, Microcomputer connection technology and so on. 2. Operation System & Network, which is a course set according to CCSE2004, mainly introduces the foundation of Operation System, Network and Communication, and focuses on Network Operation System. 3. System Analysis & Design, which introduces Object-oriental analysis, the principles and methods of design and programming, and sets up the dynamic model and static model of software system by using the unified molding language. 4. Web Development Technology, which includes two mainlines, one is ASP.NET Technology, the other is JSP/Servlet Technology, both of them should visit the database of backstage to make a complete Network application. 5. Embedded System Design, which enables students to combine the theories and laboratory content of computer software and hardware courses, and carries on teaching and experiment according to different applications. Application-oriented custom is one of the main characteristics. And other relative technologies could be revolved with the fields of Electronic Circuit, Automatic Control, Computer Principle, Computer Architecture, Assembly System, Operation System, Software Engineering and etc.

The practice proves that Suite Teaching is good, and students gave us the feedback that they could learn something practical, most important, they are proud that they can use their knowledge and abilities to fulfill the tasks which the enterprise requires.

3.3 Bilingual Teaching

To cultivate the practical and inter-disciplinary senior software talents who have international comprehensive qualities, SSNCU adopts bilingual teaching in 90% courses of Software Engineering Majors. But, please note, the Software Engineering Courses we mean don't include some computer or software technology courses, such as Discrete Mathematics, Database Architecture & Algorithm, Computer Architecture and etc. , but include the engineering management courses, such as, Introduction to Software Engineering, Software Program Management, PSP/TSP and so on. Because after continuous experiments and feedbacks, we found the technological courses are very hard for students to master by bilingual teaching, since the English level of students is so limited (not high). So if students want to obtain computer software knowledge and skills, it will be very difficult. But in the courses of Engineering Management, the knowledge points from the beginning to the end in the course are not connected very closely, the logicality is also not strong, if a student didn't master one knowledge point, he or she also can study the next knowledge very well. So even students are not good at English, they also can learn courses by Bilingual Teaching.

During the process of Bilingual Teaching, we normally employ foreign teachers to teach these courses. Foreign teachers usually change the traditional classroom teaching mode, and instruct, organize students to cooperate and exchange their thoughts between each other to cultivate students' self-thinking ability, problem-analysis ability, innovation ability and operation ability. Meanwhile, we usually emphasize on their abilities of exchange report, team cooperation and program management, and focus on cultivating students' abilities of writing English compositions, reports and graduate papers.

3.4 Workroom Training

SSNCU has supplied some doctoral science and research programs workrooms for the professors and doctors who have research projects and programs. In every workroom, there are about 10 computers and other necessary science research equipments to assist teachers to finish the science and research tasks. Until now, SSNCU has set up 10 doctoral workrooms. To strengthen students' engineering practice training and assist teachers' science and research activities, every workroom can employ some students who are good at operation to take part in the development of science and research program. Of course, since students' science and research abilities are limited, so students mainly can participate in some software design, execution and testing work. Teachers in the workroom can give students' scores as the scores of Engineering Practice Training according to students' performances. Of course, Workroom should fully prepare the relative teaching materials of engineering practice; and all materials should be approved by SSNCU Academic Committee.

Generally speaking, Students are interested in the work of software development, the realization and execution of software motivate the students' senses of achievement and improve their study interests.

3.5 Domestic and Overseas Enterprises Practice

SSNCU has set up the training-bases with famous domestic and overseas software enterprises. So far, SSNCU has set up SSNCU-enterprises training bases with more

than 20 domestic and overseas enterprises, such as ZTE Software Techonlogy (Nanchang), Jiangxi Microsoft Technology Center, An Bo Education Group Suzhou Subsidiary Company, the Training Center of Suzhou Software Garden. Every year, SSNCU organizes students to take part in the full time training programs in China or foreign countries or areas. The training time is much flexible, from one month to half a year according to different situations. Students can really experience the enterprise environment, program development environment, and directly feel the pulse of software industry through participating in the enterprises' training, operation & program development to improve their engineering abilities and professional qualities.

SSNCU has set a very strict management regulation, that is, before training, students should sign an Engineering Practice Agreement and Intellectual Property Rights and Security Agreement with enterprises and SSNCU separately, and explicit understand each party's responsibilities, and take part in the training program according to the agreements. For the overseas enterprises' training programs, SSNCU dispatches some professional teachers to cooperate with enterprises' teachers instructing students' engineering practices, so that SSNCU can know and master the students' learning effectiveness firstly and in timely. For the training programs in China, SSNCU normally doesn't need to dispatch professional teachers to manage students, but Assigns SSNCU teachers to make a reasonable training plan to supervise them, and students also have to report their work and study situations to their tutors in SSNCU regularly to ensure the smooth execution of this training plan. After finishing the training tasks, the enterprise can give the scores for students according to their performance, and even some enterprises would issue certificates of completing the relative courses for students. Besides, the instructing teachers and leading teachers of SSNCU would score and give students their scores. These two pieces of scores can decide whether the student can pass the engineering practice training or not.

4 Effectiveness of Execution

The new mode of "3+1"engineering practice teaching system has been put into practice overall in the majors teaching of Software Engineering, Computer Software, Information Security and etc., which involves about 2,000 students. Among these majors, SSNCU won the second prize of Jiangxi Provincial Teaching Achievements in 2006 because of Suite Teaching.

Although at present SSNCU has no students who graduated from SSNCU according to this new mode of "3+1"Engineering Practice Teaching System Cultivation Plan, but from the present effectiveness of teaching, the new edition of Cultivation Plan has achieved good results, and recognized extensively by teachers and students in SSNCU. Through the new mode of "3+1"engineering practice teaching, students' comprehensive qualities have improved obviously, and their operating abilities are much better than those students who haven't learnt through this mode. At present, there are some students who are participating in the domestic or overseas enterprises' training programs, and the training enterprises give us the feedback that the new mode of "3+1"Engineering Cultivation System adopted by SSNCU meets the requirement of software industry, because the students who cultivated with this mode have much stronger abilities obviously in the aspects of

software development skills, problem-resolving abilities, team cooperation, material collection, human communication and program planning. They are suitable to develop in the modern software industry much better. Students themselves believe that the first 3 years of "3+1"new mode not only enable them to understand the environment of enterprise engineering program development quickly, but also help them understand and master the theoretic knowledge of computer software; the last one year of on-site training help them adapt to the company environment very quickly and deal with the program flexibly by the enlightenment of old examples. Therefore, students can learn not only the new technologies which used by the enterprise presently, but also can adapt to the new improvement and Evolution of the new technologies in the future, and can keep the permanent competition abilities during the social competing processes in their whole lives.

Acknowledgments. This work is supported by Jiangxi Provincial Higher Education Reform Project, China (Grant No. JXJG-07-1-68).

References

1. Dai, Y., Zhang, B.: Exploration and Practice of Diverse Software talents Training Mode. Computer Education (2009)
2. Ye, L., Pan, Z.: Study on Current Guiding Platform for Project-oriented Practice Teaching. Journal of Zhejiang University of Science and Technology (2) (2009)
3. Li, H., Zhang, H., Lu, W.: Capacity-oriented Practical Teaching System in Software Engineering. Research in Higher Education of Engineering (2009)
4. Sun, X., Zhang, B.: Exploration of Practical Teaching Model in Software Engineering Based on "Project - Tutorial System". Henan Education (June 2009)
5. Lou, L.-q., Xi, B.-j.: "3+1" Teaching Mode Exploration for Textile Engineering Major. Journal of Chengdu Textile College (2) (2007)
6. Li, Y.-f.: Reflection and Exploration on the Training Pattern of Practical Talents. Journal of Hunan Radio and Television University (3) (2006)
7. Yang, X.-l.: The Study on practical Talents and the Fostering Models. Heilongjiang Researches on Higher Education (6) (2007)
8. Li, Y.: On Teachers Building to the Applied Talents in IT Undergraduates. Journal of Chengdu University (Educational Sciences Edition) (7) (2008)

Empirical Research on City Innovation Evaluation Based on Data of Some Cities in Liaoning

Yuemei Zhang and Yan Zhu

School of Business Management,
Dalian university of Technology (DLUT), China
zymay@dlut.edu.cn

Abstract. The implementation of innovation-oriented country strategy drives the urban to construct innovative cities. This paper took the latest data of six cities in Liaoning province as a sample, evaluated the innovation capability of those cities, and then put forward the corresponding policy recommendations. According to the scientific, comparability and availability principle, the index system was set up, which overcomes the problem of selecting rules unreasonable. The Entropy Weight was used to reflect the difference between objective data of indicators, which avoid inaccurate phenomenon of subjective weighting. Combining evaluation result, finally some policy was suggested, which will be helpful for evaluated object's self-management and perfect.

Keywords: innovative city, innovative city index, entropy weight.

1 Introduction

After making strategic of Building innovation-oriented country in China, there have been more than 100 domestic cities propose this ideas and set off a wave of developing an innovation-oriented city. In the beginning of 2010, National Development and Reform Commission re-deploy of Dalian, Qingdao, Xiamen, Shenyang and other cities, up to 16 cities, to be national innovation pilot cities. April 2010, issued a "guidance of Further Promoting innovation pilot cities", the Ministry of Science and Technology re-emphasized the importance of promoting the building of innovative cities, and defined the overall requirements, basic principles and priorities, at the same time, put forward specific demands of how to strengthen the organization of this work.

Liaoning province which has two national innovation pilot cities becomes one of the prominent innovation cities among provinces and provides a new opportunity for development. Particularly, existing differences between the provinces makes the evaluation of Innovation city, and then establish a benchmark and direction more necessary and urgent. This is also the reason that this paper choose to do this empirical reason.

1.1 The Meaning of City Innovation

City innovation as a new concept of urban development became vigorous after the establishment of independent innovation strategy in the country, academics pay more attention to it, but they have different opinions.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 14–23, 2011.

Shaobo Huang (2005) believes that innovation is the all-round coverage, full social participation, the whole process of linkage of the overall innovation [1]. Hui Du, Dongmei Yang, Zhihe Zhang et al (2006) supposed that the innovation city will be divided into innovation resources, innovation system and culture of innovation and other factors [2], while Fan Ye, Jiang Wei, et al (2007) also emphasized that cities' innovation based on the urban environment of innovation and innovation culture [3]. Beijing Di Fang Institute of Economic Development and Corporate Strategy (2008) consider that innovative cities is the cities which is innovation spirit prominent, rich in innovation, innovation industrial developed, and innovation environmental advantages [4].

Landry, as the founder of UK's leading research institutions COMEDIA, in his book "The Creative City", put forward the made the seven elements of dynamic mechanism of the creative city: creative people, will and leadership, Human Diversity and wisdom access, open organizational culture, a strong positive sense of local identity, urban space and facilities and access to the Internet [5].

1.2 The Review of Index System

In practice, in order to have a better understanding of city innovation, people proposed the index system of city innovation. Ling Zhu, Jin Chen, Feirong Wang (2008) designed a index system which contains three aspects —output efficiency of innovation activities, innovation resources input level, innovation and operation condition of innovation system—and total of 23 indicators [6]. Hui Du (2006) selected six basic indicators which include: the city dependence on external technology, contribution rate of technological progress to economic growth, invention patent applications accounted for the proportion of all patent applications, high-tech industrial output value of the proportion of total industrial output value, R&D funds investment proportion of GDP, business R& D input to the overall proportion of sales revenue [7]. YiShao Shi, Haiyan Bu (2008) constructed a comprehensive evaluation index system covering technology innovation, knowledge innovation, system innovation, service innovation, cultural innovation and the innovation environment, use hierarchical analysis to do comparative analysis of Innovation ability among Beijing, Shanghai and Shenzhen, which reveal the main characteristics and differences [8].

World Bank (2005) published "innovative city in East Asia" study, proposed the definition of city innovation and a series of qualitative indicators, which including: a strong research, development and innovation capacity; effective governance, efficient service; excellent transportation, telecommunications, infrastructure and functional areas and urban centers ,adequate management of culture, media, sports and academic activities Facilities; a higher level of education population; a diverse, high-quality housing options ; effective attention to environmental protection; acceptance and integration of the collision of cultures, and so on [9].

In summary, city innovation evaluation study has made some progress, from the perspective of the development of existing theory and practice, at home and abroad,

there is no unified evaluation index system and generally accepted conclusion on city innovation [10]. At least the following aspects of the deficiencies exist in present research: (1) index system is unreasonable. Multi-index system of the existing research aimed at analysis of the degree of innovation of capital cities or municipalities, but did not applicable to ground-level city. Some literature even equals the city of innovation to Science City. (2) Unreasonable evaluation index weights. Some studies using qualitative evaluation methods on the index score and the weight distribution by human caused error. (3) The lack of effective evaluation mechanisms. Existing evaluation mechanism is lack of organic connections between evaluation results and improvement measures.

In this paper, the author took the above shortcomings into consideration and insisted on the principle of scientificalness, comparability, and availability, then built a index system which contains innovation, high-tech industry and Benefits of technology. Use entropy method of objective weighting to determine the weights, and evaluated innovation of six cities in Liaoning Province comprehensively, with the data from2006 to 2008, and then put forward the associated policy recommendations.

2 City Innovation Evaluation Index System

2.1 Design Principles

(1) Scientific Principles. Scientific selection of indicators is in accordance with the essential characteristics of the city of innovation and the driving force to the selection indicators, which makes assessment results more convincing.

(2) The principle of comparability. According to different subjects, the indicator must be comparable when making comparisons, and therefore when selected the indicators this paper should follow the principle of comparability.

(3) The principle of availability. Difficulty Level of data availability has significant impact on the process of research. Considered the possibility of Indicators replaceable, and then increase the availability of indicators.

2.2 Structure of Index System

To thoroughly implement the party's Seventeenth Congress' spirit, and implement the scientific development perspective— accelerate innovation and focus on promoting economic development mode, give full play to the city in the process of building an innovative country's . Ministry of Science and Technology developed a "Creative City Development Monitoring Evaluation index (Trial)". In practice, refer to the indicators in "Creative City Development Monitoring Evaluation index (Trial)", and based on principle of scientific, comparability, availability, and the degree of difficulty in quantify indicators, remove repetitive indicators, and draw lessons from previous studies, we selected 15 high frequency indicators to evaluate the innovative capacity of cities in Liaoning Province, the specific indicators is shown in Table 1.

Table 1. The evaluation index system of cities' innovative

No.	Rule layer	Index level	unit
1	X_1 Innovation investment	X_{11} the number of technological activities in per million workers [11]	Ten thousand
2		X_{12} the number of college students in every ten thousand population [16]	Person
3		X_{13} growth rate of Fixed asset investment	%
4		X_{14} Local finance technology appropriation account for the proportion of local financial expenditure [8,13,14]	%
5	X_2 High-tech industrial	X_{21} High-tech industry added value accounts for GDP [12,15]	%
6		X_{22} High-tech industry value accounts for total industrial output value [14]	%
7		X_{23} High-tech industry added value accounts for industrial added value	%
8		X_{24} the amount of emissions of major pollutants	t
9		X_{25} the amount of decrease rate of comprehensive energy consumption in ten thousands GDP [8,13]	%
10		X_{26} Labor Productivity [12,13]	%
11	X_3 Benefic of Technology	X_{31} Internet users per million population [8,13, 14,15]	Families
12		X_{32} Urban Air Quality Index [14]	%
13		X_{33} rate of Comprehensive utilization of industrial solid waste	%
14		X_{34} Urban sewage treatment rate	%
15		X_{35} the number of Reception of foreign visitors	person

3 City Innovation Evaluation Model

Based on the above index system of evaluation index and the standardization of the raw data, this paper uses entropy method to establishment of Entropy Evaluation Model of City innovation. Empirical research was based on the data from 6 major cities of Liaoning Province from 2006 to 2008, and then policy recommendations are given.

As shown in Table 1, X_{24} is the only negative indicators; the other indicators are positive indicators. According to the scoring formula of positive indicators and Negative rate index formula, the score of the indicators can be computed.

In this paper, entropy method was used to determine the weigh. Objective weighting method based on the relationship between indicators or the variation of the index value to avoid the influence of human factors, so that the evaluation more scientific.

Let X_{ij} (i = 1,2, ..., n; j = 1,2, ..., m) be the observation data of cities i and index j, in a certain year , for a given j, the larger differences between different X_{ij} The greater the role it played in systematic comparison of the indicators. Entropy method used to determine index weight as follows, firstly, Calculate the features proportion of the system of cities i, secondly, and calculate the entropy of indexes j, lastly, calculates the entropy of each index, which are the index weights. Set W_j for

the Entropy evaluation indexes j, Let Pi be the comprehensive evaluation score of cities i in a certain year, based on the linear weighted formula, the city innovation evaluation scores can be calculate as follows:

$$p_i = \sum_{j=1}^{m} w_j p_{ij} \quad (1)$$

4 Samples and Index Score

This study selected six cities 'data in Liaoning Province from 2006 to 2008 as samples. Index is expressed as X_{ij}, where i = Shenyang, Dalian, Fushun, Benxi, Dandong, Jinzhou; j = 1,2, ..., 15, The majority raw data of index is from the Statistical Yearbook in year 2007,2008,2009, and China Economic Information Network, as well as local sites. Original data is shown in table 2 below. Using the method shown above, the value of six cities related indicators, Entropy and Entropy weight, the results of the evaluation can be calculated, and the results of the relative speed evaluation was shown in table 3.

4.1 Analysis of the Result

(1) Comprehensive innovation evaluation of Dalian and Shenyang is far ahead

Making use of Evaluation results in Table 6, take the year 2006 to 2008 as the abscissa and the city's comprehensive evaluation scores as the vertical axis, we can get Figure 1.

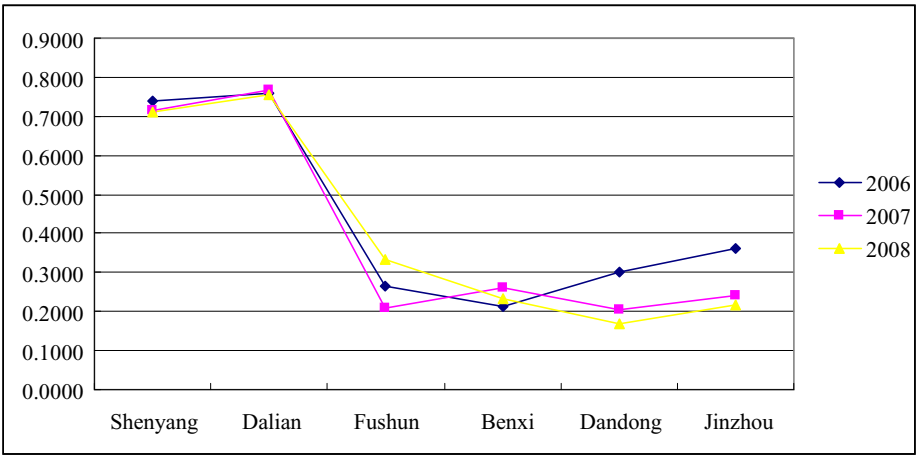


Fig. 1. The carve of Liaoning 2006-2008 create city comprehensive evaluation

From Figure 1, we can clearly see that innovation of Dalian and Shenyang has been leading in the province in year 2007, 2008, 2009. The result is consistent with the country's decision that choose Dalian and Shenyang as the innovation pilot cities,

on the other hand, this illustrate the appropriate of indicators and methods which was chose in this paper. The relative ranking of innovation in other cities appeared relatively large fluctuations, while Dandong, Jinzhou relative ranking dropped, Fushun relative ranking rose.

(2) The innovation growth rate of Dalian and Shenyang is not the first

From the relative quantitative relationship between cities, which shown in Table 3, the speed of innovation can be: In 2007, the city ranked as follows: Benxi> Dalian> Shenyang> Fushun> Dandong> Jinzhou; in 2008, the city ranked as follows: Fushun > Shenyang> Dalian> Jinzhou> Benxi> Dandong. That is to say, from the aspect of incremental innovation, Benxi and Fushun in 2007 and 2008, respectively, had faster growth than that of Shenyang and Dalian. However, the volatility of results of comprehensive evaluation of Benxi and Fushun is also great, not stable. As for Dalian and Shenyang, the relative growth rate of innovation, though not the first, the results of comprehensive evaluation of its city innovation is relatively stable.

(3) Factors which Affect City innovation

In each year, which indexes are heavily weighted are as follows: X_{12}, X_{14}, X_{26}, X_{34}, X_{35}, so that these indicators constitute the main factors affecting the city's innovation.

In 2006, the indicator X_{35} become the biggest weight, which indicates that the number of foreign exchange greatly affect the results of the evaluation of city innovation. Dalian, with beautiful landscape and the position of the opening window, and Shenyang, as the capital city, with a long history, attracting a large number of foreign tourists, all this contributed a great deal of innovation and strength to city innovation. In 2007, Fushun has a rapid growth of indicators X_{35}, while the pace of city innovation has been enhanced. In the subsequent 2007 and 2008, the weight of this indicator X_{35} decreased which means that the external forces are weakening in city innovation, while enhance the strength of the city itself be more emphasized. Thus, X_{14}, X_{26}, X_{12} and X_{34}whose weight is around 0.1 shows the internal factors of the city's innovation.

4.2 Some Policy Proposals

The analysis above shows that the innovative capacity of Dalian and Shenyang has been in a leading state.

In 2006, the State Council brought "revitalize the old industrial bases in Northeast China" into the "Eleventh Five-Year" plan and officially approved in August 2007. Under the guidance of national policy, the government's strong support and the tireless efforts of all departments, there is a significant increase of city innovation capacity in Liaoning Province and "Revitalizing the Northeast Plan" is also beginning to be seen. In the "Northeast revitalization plan" proposed an exciting vision that the Northeast region become into a higher level of overall economic development after 10 to 15 years' efforts, which is to be expected.

In the past few years, economic development of Liaoning Province has been rapidly increased, and achieved certain results; however, the establishment of innovation-oriented city is a complicated systematic project, which including a multi-dimensional factors. So this paper made the following suggestions:

(1) Enhance independent innovation capability. Based on the data above, we can see that the development of high-tech industry contribution little to the city's innovation. The Government should lead business to build innovation system and then form a platform for the transformation of science and technology. The significantly increase of local government technology funding of Shenyang and Dalian in 2007 shows that the government has a better understanding of the technology economy which is conducive to the economic restructuring.

(2) Enhance environment comfort of the city. Environmental comfort refers to not deliberately take any winter protection or summer cooling measures and the degree whether people feel comfortable under the natural environment. The improvement of environmental comfort is a policy of benefiting, and urban air quality is to be improved. Therefore, in order to development a "low-carbon economy", energy conservation is one effective way and was brought into "second Five-Year" plan. .Low-carbon energy is the basic guarantee for low-carbon economy; what's more, cleaner production is the key to low-carbon economy, recycling is an effective method for low-carbon economy and sustainable development is the fundamental direction of low-carbon economy. Therefore, transforming economic growth, promote economic structural adjustment, saving resources and energy are the responsibility of each city.

5 Conclusion

This paper is based on the principle of scientific, comparability, availability, through there aspect rule layer—innovation investment, and high-tech industry and benefit of technology,establish evaluation index system of city innovation, then use entropy method to determine the weights of each index, and comprehensive evaluated city innovation in Liaoning Province, whose sample included six major cities and the data is from 2006 to2008 .in the end of this paper, policy recommendations which associated with the evaluation results was proposed. Features and innovations of this article are as follows:

①The sample of existing evaluation of city innovation is limited to provincial capitals and sub-provincial city, according to the principles of scientific, comparability and accessibility, this paper establishes the general index system of the city's innovation and expanded application of the Index System; ② using entropy weight method to perform multi-year and multi-city assessment reflected difference between objective data of each indicator and the subjective weighting method avoid the problem which caused by man-made weight distribution , received evaluation results which was conduct in the same year; ③ combining the evaluation results, this paper put forward targeted policy recommendations, and found the organic connection between establishment of the evaluation results and improvement measures, which is conducive to the object's self-management evaluation and improvement.

Table 2. Initial data

Rule layer	Index Level	Shenyang			Dalian			Fushun			Benxi			Dandong			Jinzhou		
		2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008	2006	2007	2008
X_1	X_{11}	264.25	291.63	339.67	250.95	213.10	182.67	175.75	105.53	117.67	275.76	166.49	171.51	115.44	72.83	84.59	205.51	106.15	113.86
	X_{12}	648.20	447.30	458.40	401.00	434.00	450.00	127.00	142.00	159.00	86.40	87.20	88.00	82.49	86.94	103.79	214.00	229.00	247.00
	X_{13}	31.30	30.49	25.38	32.30	31.40	36.04	25.96	40.20	37.30	6.50	9.70	36.11	30.50	45.39	32.19	30.40	55.60	45.80
	X_{14}	0.47	3.17	3.43	0.22	3.95	4.12	0.38	0.91	0.94	0.27	0.91	0.95	1.10	1.72	1.45	1.13	1.13	1.09
X_2	X_{21}	13.16	15.86	16.58	16.34	19.23	15.68	12.04	12.79	13.09	7.70	9.30	9.30	4.78	6.27	7.20	8.03	8.07	8.47
	X_{22}	49.72	47.21	48.10	57.65	67.72	76.93	32.30	35.93	45.53	17.10	20.65	18.91	60.85	61.77	64.33	28.48	29.25	32.59
	X_{23}	41.08	38.21	50.21	56.13	52.03	43.27	46.74	60.03	76.79	43.02	55.80	17.39	75.67	83.15	26.17	35.04	47.52	27.04
	X_{24}	68468	92684	92665	98207	95172	101242	81467	90494	7415	136900	12355	103330	36800	39629	43546	66064	65296	70990
	X_{25}	5.40	4.50	6.01	4.40	4.00	5.70	5.00	4.60	4.80	4.40	4.40	4.40	8.40	8.40	8.40	5.30	5.00	5.30
	X_{26}	10.45	12.92	12.32	9.80	6.60	5.90	4.70	2.62	2.83	3.32	2.26	6.10	1.99	1.52	1.86	6.68	2.88	2.89
X_3	X_{31}	1141.65	1374.80	1464.18	1080.83	1287.41	1445.20	436.10	753.06	888.80	684.98	686.29	957.35	250.86	359.03	516.85	390.60	562.54	640.72
	X_{32}	87.95	88.49	88.50	92.60	93.89	96.40	80.55	82.19	84.93	79.45	86.58	91.78	94.80	92.30	83.00	83.20	76.30	94.30
	X_{33}	91.92	91.16	92.29	78.69	91.07	95.73	65.41	65.26	56.83	39.71	37.65	38.39	62.00	57.28	49.11	49.46	46.30	46.30
	X_{34}	73.00	72.98	76.60	80.37	81.08	87.18	54.10	48.78	56.88	76.16	82.73	83.55	0.00[1]	0.00	0.00	36.24	43.13	51.50
	X_{35}	400037	456013	476357	700032	840031	950045	33097	60869	77201	82199	135593	217816	128000	148524	190762	68117	87342	114820

[1] Note: As the original data of Dandong in the table above, the sewage treatment rate is zero. The author is also filled with doubts, but the data does not affect the evaluation results. The data of Index X_{34} of Dandong is from Statistical Yearbook of Liaoning Province (in year 2007), which is in the second part of the "basic statistics provincial cities" in the "Environmental Protection (second)" data sheet. Similarly, in year 2007 and 2008, the data of X_{34} of Dandong is from the Statistical Yearbook of Liaoning Province (2008, 2009), which is in the second part of the "basic statistics provincial cities" in the "Environmental Protection (second)" data sheet.

Table 3. The results of the relative speed evaluation

period	In year 2006						In year 2007						In year 2008					
city	Shenyang	Dalian	Fushun	Benxi	Dandong	Jinzhou	Shenyang	Dalian	Fushun	Benxi	Dandong	Jinzhou	Shenyang	Dalian	Fushun	Benxi	Dandong	Jinzhou
Comprehensive score	0.7390	0.7614	0.2633	0.2144	0.3001	0.3626	0.7141	0.7678	0.2097	0.2610	0.2058	0.2405	0.7101	0.7541	0.3334	0.2316	0.1697	0.2173
Relative score	1.0000	1.0303	0.3564	0.2902	0.4061	0.4906	1.0000	1.0753	0.2936	0.3655	0.2882	0.3368	1.0000	1.0619	0.4695	0.3262	0.2390	0.3060
		1.0000	0.3459	0.2817	0.3942	0.4762		1.0000	0.2730	0.3399	0.2680	0.3132		1.0000	0.4421	0.3072	0.2251	0.2882
			1.0000	0.8143	1.1395	1.3765			1.0000	1.2449	0.9816	1.1471			1.0000	0.6948	0.5091	0.6518
				1.0000	1.3994	1.6906				1.0000	0.7885	0.9215				1.0000	0.7327	0.9381
					1.0000	1.2081					1.0000	1.1686					1.0000	1.2803

Acknowledges

This work was supported by Economic and social development issues in Liaoning province, under the Grants No: 2010lslktglx-48.

References

1. Huang, S.: The ideas and construction of city Innovative. Journal of Guilin University of Electronic Technology 25(6), 97–100 (2005) (in Chinese)
2. Yang, D., Zhao, L., Yan, L.: Innovative City: A conceptual model and development patterns. Journal of Scientology and Technology Management 8, 97–101 (2006) (in Chinese)
3. Ye, F.: Innovative city construction elements and implementation path of Innovative city. Journal of Shenyang Cadres 22(4), 35–37 (2007) (in Chinese)
4. Beijing Di Fang Institute of Economic Development and Corporate Strategy, "connotation, influencing factors and developmental stage of the city Innovation". In: Proceedings of Urban Development and Innovation Symposium, vol. (5), pp. 1–5 (2008) (in Chinese)
5. Landry, C.: The Creative City: A Toolkit for Urban Innovators, 1st edn. Earth scan Publications Ltd., London (2000)
6. Zhu, L., Chen, J., Wang, F.: Study of evaluation system of city innovation development. Scientific Research 2, 215–222 (2008) (in Chinese)
7. Du, H.: Connotation and characteristics of Innovative cities. Journal of Dalian Official 22(2), 10–12 (2006) (in Chinese)
8. Shi, Y., Bu, H.: Evaluation system of Innovative city and its comparative analysis. China Science and Technology Forum 1, 22–26 (2008) (in Chinese)
9. World Bank, Innovative city in East Asia research report (2005)
10. Zhang, J., Liu, Y.: Research result and shortage of Innovative city construction. Journal of Modern Economics and Management 8, 81–83 (2009) (in Chinese)
11. Wei, J., Liu, Y., Hu, S.: Evaluation research of innovative city,based on principal component analysis. Journal of Hunan University (Social Science Edition) 23(3), 53–58 (2009) (in Chinese)
12. Zhou, N.: Evaluation system and assessment method of Innovative city construction. Journal of Statistics and Decision 8, 21–23 (2008) (in Chinese)
13. Wang, R., Deng, P.: The construction of Innovative city evaluation index system. Journal of Industrial Technology and Economy 27(1), 69–73 (2008) (in Chinese)
14. Liu, R., Qiu, D.: Empirical study of Innovative city evaluation index system - Based on correlation analysis. Journal of Technical and Economic 29(1), 29–34 (2010) (in Chinese)
15. Yang, Z., Yu, J., He, L.: Design of evaluation index system of the Innovation city. Theory Monthly 5, 151–153 (2009) (in Chinese)
16. Wei, J., Yi, L., Hu, S.: Development Strategy of Creative City in Hangzhou. Journal of Hangzhou Science and Technology 3, 34–36 (2007) (in Chinese)

Video Vehicle Detection and Tracking System

Qinjun Li, Hongshe Dang, Yujie Zhang, and Danhui Wu

College of Electric and Information Engineering Shaanxi University of Science and Technology, Xi' an, China
liqinjun@sust.edu.cn

Abstract. Non-intrusive video vehicle detection and tracking for traffic flow surveillance and statistics is the primary alternative to conventional inductive loop detectors. Vision-based systems for traffic have an impressive spread both for their practical application and interest as research issue. This paper presents vision-based vehicle detection and tracking system which consists of environment background segmentation and subtraction, foreground moving object extraction, moving vehicles detection algorithms, object tracking algorithms, and vehicle classification. The proposed system can perform well for the video sequences acquired under different weather, illumination, and traffic conditions through the use of these technologies.

Keywords: video; vehicle detection; tracking; vehicle classification; ITS.

1 Introduction

In order to solve the increasingly serious traffic problems, the integrated use of computers, databases and image processing and other high-tech intelligent transport system (ITS) came into being, the vehicle video detection technology is also more and more attention of scholars home and abroad. Video vehicle detection is the rise vehicle detection technology in recent years, compare with the traditional ring coil detection technology, microwave detection technology, it has a number of additional advantages, including the simple installation and maintenance, rich information acquisition and wide application, low cost and so on. This technology for a traffic measurement system that collects traffic data from roadways and hinges plays a critical role in many ITS applications, such as adaptive traffic signal control, path planning, and traveler information services. Apart from real-time operations, traffic data collected over time can also serve as an important resource for long-term planning and administrative activities. Video vehicle detection technology available including gray comparison, frame difference, background subtraction, optical flow and road markings laws.

2 Vision-Based Traffic Measurement System Overview

The vision-based traffic measurement system consists of a stationary video camcorder, a host computer and a number of communication devices. The camcorder

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 24–31, 2011.

continuously captured video sequences from a roadway. The collected sequences are immediately transmitted through the communication equipment to the host computer installed in a traffic information center. The video sequences are processed and interpreted in the host computer.

The camcorder can be installed on an overpass, a tall building or a high pole. Fig.1 depicts a configuration of the camcorder. Several configuration parameters regarding the camcorder are given in the figure, including the depression (i.e. tilt) angle θ, the roll angle γ, the viewing (i.e. pan) angle λ, and the height *h* of the camcorder. We assume that the roll angle γ equals 0. The parameters θ, *h* and the focal length *f* of the camcorder are known beforehand. As for the viewing angle λ, it will be determined during camcorder calibration.

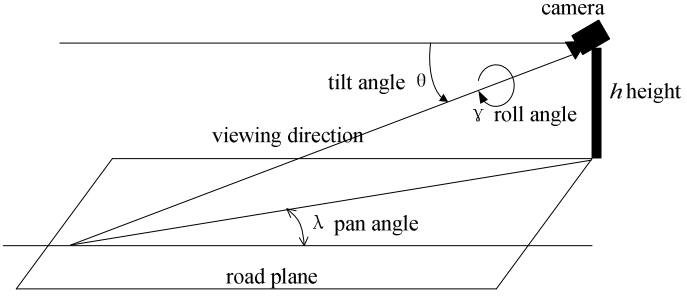


Fig. 1. Configuration of the camcorder

The proposed system can be applied to the public traffic images which are similar to those taken by the public camcorders that have been widely installed by our Traffic Administration Bureau along freeways, highways and main streets. A block diagram for the proposed system is shown in Fig.2. Four major components constitute the diagram, they are, and foreground moving objects extraction, vehicle detection, vehicle tracking, and vehicle classification.

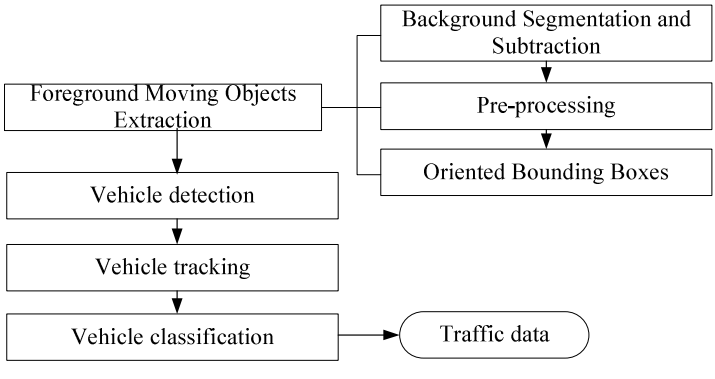


Fig. 2. Block diagram of the proposed vision-based traffic measurement system

3 Foreground Moving Object Extraction

Tracking in outdoor, crowded scenes requires that the tracked entities can be segmented out reliably in spite of the complexities of the scene due to changing illumination, static and moving shadows, uninteresting background and camera motion. The method should also be fast enough so that no frames are skipped. Another requirement in this application is that stopped entities such as vehicles or pedestrians waiting for a traffic light should continue to be tracked.

3.1 Background Segmentation and Subtraction

An adaptive Gaussian mixture model method is used. Each pixel in the image is associated with a mixture of Gaussian distributions (5 or 6) based on its intensities. Each distribution is characterized by a mean and variance and a weight representative of the frequency of occurrence of the distribution. The Gaussian distributions are sorted in the order of the most common to the least common distribution and the pixels with matching distribution having a weight above a certain threshold are classified as background while the rest are classified as the foreground.

Moving entities are then extracted using a two pass connected components extraction method. In order to eliminate noise from being classified as foreground, a threshold is used so that any blob with area lower than the threshold is deleted from the foreground.

To segment the vehicle objects form a complex background, we use the well-known background subtraction method. Here we use a self-adaptive background subtraction method. The basic principal of our method is to modify the background image (called the current background (B)) by using instantaneous background (IB) and applying an appropriate weighting α as follows $CB_{k+1}=(1-\alpha)CB_k+\alpha IB_{k+1}$ where the subscript k is the frame number index, the instantaneous background is defined as $IB_k= M_k \bullet CB_k+(\sim M_k)\bullet I_k$, I_k is the current image frame, M_k is the binary vehicle mask. It is obtained by subtracting the brightness of the current image from the estimated current stationary background followed by a binary thresholding.

3.2 Pre-processing

After background subtraction, we may find some noise of the extracted foreground image. Here, we develop the pre-processing process to obtain a more clean-cut silhouette for the following vehicle objects tracking. There are four steps in the pre-processing: noise removal, morphological filtering, labeling, and size filtering.

3.3 Oriented Bounding Boxes

Horizontal and vertical axis aligned boxes cannot provide tight fits to all vehicles moving in arbitrary directions. As a result, oriented bounding boxes are used to represent blobs. The oriented bounding box is computed using the two principal axes of the blob, which are in turn computed using principal component analysis. The covariance matrix used to compute this consists of the blob's first and second-order moments

$$\begin{bmatrix} M_{20} & M_{11} \\ M_{11} & M_{02} \end{bmatrix} \tag{1}$$

Where, M_{ij} is the $(i,j)^{th}$ order moment of the blob. Diagonalizing M gives

$$M=\Delta^{T}D\Delta\ . \tag{2}$$

where$\Delta=[v_1\ v_2]$represents the eigenvectors and

$$D=\begin{bmatrix} e_1 & 0 \\ 0 & e_2 \end{bmatrix} \tag{3}$$

represents the eigenvalues. If $e_1>e_2$, we choose v_1 as the principal axis with elongation $2\cdot e_1$. The angle made by the principal axis with respect to the x axis of the image is also computed from the eigenvectors. Similarly, v_2 is chosen as the second principal axis with elongation $2\cdot e_2$.

4 Moving Vehicles Detection Algorithms

In this section, we outline the algorithms used to perform vehicles segmentation in detail, pointing out the day-time condition. To explain the double-difference operator, we can consider the sequence of binary images $\{I_m\}$: the difference-image Dm is defined as:

$$D_n(i,j)=\left|I_n(i,j)-I_{n-1}(i,j)\right| \tag{4}$$

The double-difference image is obtained by performing a logical and between pixels belonging to two subsequent difference-images, thresholded by a threshold T:

$$DD_n(i,j)=\begin{cases} 1 & if\ (D_{n+1}(i,j)>T)\wedge(D_n(i,j)>T) \\ 0 & otherwise \end{cases} \tag{5}$$

We observed that this operator is quite immune to noise, due to the non-repeatability of noise in three subsequent frames. Moreover, the double-difference image detects motion on the central frame In, where the image luminance gradient will be computed.

VTTS performs motion detection and luminance variation detection in parallel. In order to obtain the spatial-temporal extraction of the targets, both of these data have to be joined. We propose a new operator called *MEC* (*Moving Edge Closure*): a pixel of the image is labelled as moving if its gradient magnitude (∇I_n) is high enough and if a pixel previously labelled as *moving* belongs to the pixel's neighbourhood :

$$MEC_n^0(i,j)=\begin{cases} 1\ if\ DD_n(i,j)=1 \\ 0\ otherwise \end{cases} \tag{6}$$

$$MEC_n^r(i,j) = \begin{cases} 1 \; if \; MEC_n^{r-1}(i,j) = 1 \vee (\nabla I_n(i,j) \geq T_G) \\ \qquad\qquad \wedge \exists (k,l) \in X \left| MEC_n^{r-1}(k,l) = 1 \right. \\ 0 \; otherwise \end{cases} \tag{7}$$

where T_G is the value of the threshold applied onto the gradient image and X is the pixel's neighbourhood.

Nevertheless, an approximation of the MEC operator suited for hardware implementation is required. In fact, as shown in the definition of the operator (reported above), MEC needs a recursive algorithm to be performed. As a matter of fact, raster-scan algorithm able to achieve realtime behaviour can not fulfil recursive algorithm. Therefore, MEC algorithm has been partitioned in two steps: direct scan and reverse scan. In the former, the image is analyzed from the upper left corner to the lower right, while in the latter from the lower right to the upper left. This approximation is effective, except when a strongly concave object is present: in such a case some contours of the moving object could not be caught. Nevertheless, this is quite a rare situation in real image sequences under commonly used points of view.

A further drawback of the approximation introduced is the wrong link established between close objects connected by a stationary area with high gradient (such as shadows). To overcome this, MEC approximation has to be reviewed.

5 Kalman Filter Tracking

The results from this module are passed onto the high-level module where tracking consists of refining the position and shape measurements by means of Kalman filtering. An Extended Kalman filter is used for estimating the position of moving object in scene coordinates while shape of the moving object is estimated in image coordinates using a discrete Kalman filter.

The position estimation filter is responsible for estimating the target position in scene coordinates. The entities are assumed to move with constant velocities and any changes in the velocity are modeled as noise in the system. Because of the nonlinearity in the mapping from the state space (world coordinates) to the measurement space (image coordinates), an extended Kalman filter is used. The state vector is represented as $X = [x, \dot{x}, y, \dot{y}]$, where x, y are the positions of the centroid in the *x*-*y* scene coordinates and $\dot{x}, \dot{y}$ are the velocities in the x, y directions. The state transition matrix is given by

$$\begin{bmatrix} 1 & \delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \delta t \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

where δt is the time elapsed between two frames. The error covariance of the system noise is given by

$$Q = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} q \tag{9}$$

Where

$$A = \begin{bmatrix} \frac{(\delta t)^3}{3} & \frac{(\delta t)^2}{2} \\ \frac{(\delta t)^2}{2} & \delta t \end{bmatrix} \tag{10}$$

and q is the variance in the acceleration.

The measurement error covariance R_k is given by

$$\begin{bmatrix} \sigma_k^2 & 0 \\ 0 & \sigma_k^2 \end{bmatrix} \tag{11}$$

The measurement error standard deviation is obtained based on the variance in the percentage difference in the measured and previously estimated size (area). The Jacobian of the measurement matrix H is used due to the nonlinearity in the mapping from image to world coordinates of the target's positions.

The filter is initialized with the scene coordinate position of the object obtained by back projecting the image measurements using the homography matrix. The homography matrix is computed from the camera calibration. The filter estimates a model of the motion of the target based on the measurements. The estimate of the model corresponds to estimating the position and the velocity of the target.

6 Vehicle Classification

Our system classifies as many as seven types of vehicles. Here, we develop a so-called hierarchical vehicle classification method. The hierarchical classification consists of coarse classification and fine classification. The seven identifiable vehicle types are: (A) sedan, (B) van, (C) pickup, (D) truck, (E) van truck, (F) bus, and (G) trailer, as shown in Fig.3.

Fig. 3. Seven vehicle types

Based on the silhouette of the vehicles, we can extract three different features as: length, aspect ratio, and compact ratio. The length of the vehicle is the dominant feature which can be used to do the coarse classification to classify the vehicle in to a large or a small one. Then we may use other features to do the fine classification, which takes advantage of the second and the third feature. The classification process is shown as follows:

1. Coarse classification identifies the moving object as a large vehicle or a small vehicle.
2. If it is a small vehicle, then calculate the following $F_1 = \alpha_1 \bullet$ (aspect ratio) + $\beta_1 \bullet$(compact ratio) where $\alpha_1=30$ and $\beta_1=1$. If $F_1>$THHC, then it is a sedan else it can be a van or pick-up.
3. To differentiate the van from the pick-up truck, we examine the height/length ratio (HLR). Since the pickup-truck with framed loading compartment will be higher, we may compare the HLR to find the pickup-truck with framed loading compartment.
4. Differentiate the pickup-truck from the van by comparing the van their compact ratio (CR).
5. Fine classification divides the larger vehicles into two groups: (a) bus and van truck; (b) truck and trailer. The first group has a larger CR than the second group.
6. To separate bus from van truck, compare their HLR.
7. Calculate the following measurement to separate the trailer from the truck as $F_2 = \alpha_2 \bullet$(length)+ $\beta_2 \bullet$(area/length), where $\alpha_1=1$ and $\beta_2=2$. If $F_2>$THLC, then it is a trailer else it is a truck.

7 Conclusion

We have presented a vision-based traffic measurement system—one of the most important components of any ITS. The system consisting of a CCD camera, a host computer and a number of communication devices is easy to move, install and operate. The proposed system adopted several intriguing techniques such as background segmentation and subtraction, moving vehicles detection, object tracking, vehicle classification and so on, to make the traffic data acquire by our system accurately and efficiently.

Acknowledgments. This paper Supported by The Xianyang Municipal Planning Project of Science and Technology (2010K08-07); Natural Science Foundation of Shaanxi University of Science and Technology (ZX09-33).

References

1. Veeraraghavan, H., Masoud, O., Papanikolopoulos, N.P.: Senior Member, IEEE: Computer Vision Algorithms for Intersection Monitoring. IEEE Transactions On Intelligent Transportation Systems 4(2) (June 2003)
2. Baker, K.D., Sullivan, G.D.: Performance assessment of modelbased tracking. In: Proc. IEEE Workshop Applications of Computer Vision, Palm Springs, CA, pp. 28–35 (1992)

3. Dailey, D.J., Cathey, F.W., Pumrin, S.: An Algorithm to Estimate Mean Traffic Speed Using Un-Calibrated Cameras. IEEE Trans. on Intelligent Transportation Systems 1(2), 98–107 (2000)
4. Klein, L.A.: Sensor Technologies and Data Requirement for Intelligent Transportation Systems, pp. 69–75. Artech House, Boston (2001)
5. Coifman, B., et al.: A Real-Time Computer Vision System for Vehicle Tracking and Traffic Surveillance. Transportation Research-c, 102–115 (December 1, 1998)
6. Veeraraghavan, H., Masoud, O., Papanikolopoulos, N.P.: Vision-based monitoring of intersections. In: Proc. IEEE Conf. Intelligent Transportation Systems, pp. 216–225 (2002)
7. Wang, J.M., Chung, Y.C., Chang, S.L., Chen, S.W.: Lane Marks Detection Using Steerable Filters. In: Proc. of 16th IPPR Conf. on Computer Vision, Graphics and Image Processing, pp. 858–865 (2003)
8. Lai, A., Yung, N.: Vehicle-Type Identification Through Automated Virtual Loop Assignment and Block-Based Direction-Biased Motion Estimation. IEEE Trans. on ITS (2), 86–97 (2000)
9. Welch, G., Bishop, G.: An introduction to the Kalman filter (tutorial). In: Proc. SIGGRAPH, pp. 23–30 (2001)

An Image Encryption Scheme with Efficient Permutation and Diffusion Processes

Ruisong Ye

Department of Mathematics, Shantou University
Shantou, Guangdong, 515063, P.R. China
rsye@stu.edu.cn

Abstract. This paper proposes an efficient chaos-based image encryption scheme, in which shuffling the positions of image pixels incorporates with changing the grey values of image pixels to confuse the relationship between the cipher-image and the plain-image. In both the permutation process and the diffusion process, generalized Bernoulli shift maps are utilized to generate chaotic orbits, which are used to get an index sequence for the permutation of image pixel positions and to yield two random grey value sequences for the diffusion of grey values, respectively. Experimental results are carried out with detailed analysis to demonstrate that the proposed image encryption scheme possesses large key space to resist brute-force attack and possesses good statistical properties to frustrate differential attacks as well.

Keywords: generalized Bernoulli shift map; chaotic system; image encryption.

1 Introduction

Chaos has been introduced to cryptography as its ergodicity, pseudo-randomness and sensitivity to initial conditions and control parameters are close to confusion and diffusion in cryptography. These properties make chaotic systems a potential choice for constructing cryptosystems. In the past decades, digital image information has been widely communicated over the Internet and wireless networks thanks to the rapid developments in the multimedia industry and communications. Therefore the security issue of digital images has attracted more attentions. Due to some intrinsic features of digital images, such as bulk data capacity and high correlation among adjacent pixels, traditional encryption algorithms, such as DES, RSA, are thereby not suitable for practical digital image encryption. Fortunately, chaos-based image encryption algorithms have shown their superior performance[1–4]. Chaos-based image encryption schemes are usually composed of two processes generally: chaotic confusion of pixel positions by permutation processes and diffusion of pixel grey values by diffusion functions, where the former permutes a plain-image with chaotic systems, and the latter changes the value of each pixel. A good permutation process should show good shuffling effect and a good diffusion process should cause great modification over the cipher-image even if only a minor change for one pixel in the plain-image.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 32–39, 2011.

Recently, some chaos-based image encryption algorithms are broken due to their small key spaces and weakly secure encryption mechanism[5, 6]. To overcome the drawbacks such as small key space and weak security in chaos-based image encryption algorithms, many researchers turn to find some improved chaos-based cryptosystems with large key space and good diffusion mechanism[3, 7]. In this paper, an efficient image encryption scheme with a permutation-diffusion structure is proposed. In both the permutation process and the diffusion process, generalized Bernoulli shift maps are utilized. First, the permutation process employs one generalized Bernoulli shift map to generate a pseudo-random real number sequence, which is then sorted and applied to permute the image pixel positions totally. To improve the diffusion effect, a two-way diffusion process is presented, where another two generalized Bernoulli shift maps are utilized to generate two pseudo-random grey-scale value sequences. The two sequences are then used to modify the pixel grey values sequentially. The yielded grey-scale value sequences are not only sensitive to the control parameters and initial conditions of generalized Bernoulli shift maps, but also strongly depend on the plain-image processed, therefore the proposed scheme can resist differential attack, known-plaintext attack as well as chosen-plaintext attack. The proposed scheme also possesses large key space, therefore frustrating brute-force attack.

2 The Generalized Bernoulli Shift Map

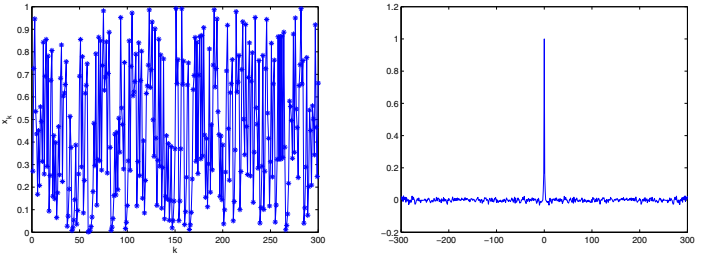
The Bernoulli shift map $B_0 : [0, 1] \to [0, 1]$ is given by

$$x_{n+1} = B_0(x_n) := 2x_n \bmod 1 = \begin{cases} 2x_n, & \text{if } x_n \in [0, 1/2) \\ 2x_n - 1, & \text{if } x_n \in [1/2, 1] \end{cases} \tag{1}$$

The Bernoulli shift map yields a simple example for an essentially nonlinear stretch-and-cut mechanism, as it typically generates deterministic chaos. Such basic mechanisms are also encountered in more realistic dynamical systems. Note that 'stretch and fold' or 'stretch, twist and fold' provide alternative mechanisms for generating chaotic behavior. In this paper, we shall consider its generalized version shown as

$$x_{n+1} = B(x_n) := \frac{x_n}{a} \bmod 1. \tag{2}$$

where $x_n, x_{n+1} \in [0, 1]$ are the states of the map, and $a \in (0, 1)$ is the control parameter. As $a = 0.5$, B becomes the regular Bernoulli shift map (1). A typical orbit of x_0 derived from the dynamical system is $\{x_k = B^k(x_0), k = 0, 1, \cdots\}$, which is shown in Fig. 1(a) for $a = 0.3731, x_0 = 0.2709$. Its waveform is quite irregular and indicates that the system is chaotic. The control parameter a and the initial condition x_0 can be regarded as cipher keys as the map is used to design image encryption schemes. There exists some good dynamical features in generalized Bernoulli shift maps, such as desirable auto-correlation and cross-correlation features.



(a) A typical orbit of generalized Bernoulli shift map

(b) Auto-correlation

Fig. 1. A typical orbit of the generalized Bernoulli shift map with $a = 0.3731, x_0 = 0.2709$

3 The Image Encryption Scheme Based on Generalized Bernoulli Shift Maps

3.1 Permutation Process

In this subsection, we propose a permutation process to confuse plain-image totally. Thanks to the chaotic nature of generalized Bernoulli shift maps on the unit interval $[0, 1]$, one can easily get the chaotic orbit $\{x_k, k = 0, 1, \cdots\}$ of x_0 with given control parameter a. We rearrange all the x_k values of the orbit to get one new sequence $\{\bar{x}_k, k = 0, 1, \cdots\}$ according to the order from small to large. As a result, we also get an index order number for every x_k. The index order number sequence can be applied to permute the image pixel positions and therefore confuse the image to get a shuffled image. The permutation process is stated as follows.

Step 1. Set the values of the control parameter a and the initial condition x_0.

Step 2. Iterate the generalized Bernoulli shift map (2) to get the truncated orbit of x_0, say $\{x_k, k = 0, 1, \cdots, H \times W\}$ where H, W are the height and the width of the processed image respectively.

Step 3. Sort $\{x_k, k = 1, \cdots, H \times W\}$ to get an index order sequence $\{Ix_k, k = 1, \cdots, H \times W\}$.

Step 4. Reshape the grey-scale value matrix of the processed plain-image A sized $H \times W$ to one vector U with length $H \times W$; permute the vector U by Ix in the following way to get one new vector V: $V_k = U_{Ix_k}, k = 1, \cdots, H \times W$.

Step 5. Reshape V back to one 2D matrix to yield the shuffled image B.

3.2 Diffusion Process

The diffusion process will enhance the resistance to statistical attack and differential attack greatly, in which the histogram of the cipher-image is fairly uniform and is significantly different from that of the plain-image. The opponent can not find any useful clues between the plain-image and the cipher-image and so can

not break the cryptosystem even after they have spent a lot of time and effort. A good diffusion process should use a keystream strongly related to plain-image. When encrypting different plain-images (even with the same cipher keys), the encryption scheme should generate different keystreams. The diffusion process is outlined as follows.

Step 1. Applying the permutation process to confuse the plain-image A and get a shuffled image B. Set the values of the initial conditions y_0, z_0 and the control parameters b, c, say $y_0 = 0.7627, b = 0.3216, z_0 = 0.3607, c = 0.3902$, in the diffusion process.

Step 2. Let $i = 0$.

Step 3. Apply the following quantization formula to yield a 8-bit random grey value $d(i)$:

$$d(i) = \text{floor}(L \times y_i).$$

where L is the color level (for a 256 grey-scale image, $L = 256$), the "floor" operation on x returns the largest value not greater than x.

Step 4. Compute the pixel grey value in the cipher-image by a two-point diffusion transmission:

$$C(i+1) = \phi(i+1) \oplus [(d(i) + C(i)) \text{mod } L], \tag{3}$$

where $\phi(i+1)$ is the grey value of the current operated pixel in the shuffled image which has been rearranged according to the order of row or column to a vector with length $H \times W$, $C(i)$ is the previous output cipher-pixel grey value. The diffusion process is well defined as the initial condition $C(0)$ is provided. $C(0)$ can be set to be part of the keys in the diffusion process or can just take the value of $d(0)$ for simplicity.

Step 5. Compute s by $s = 1 + [C(i+1) \text{mod } 2]$ to get the next y_{i+1} by iterating the generalized Bernoulli shift map $y_{i+1} = B^s(y_i)$ with control parameter b. This is the crucial step to generate a keystream depending on the plain-image since s is related to $C(i+1)$, so is y_{i+1}. The encrypted image not only relates to the cipher keys, but also relates to the plain-image.

Step 6. Let $i = i + 1$ and return to Step 3 until i reaches $H \times W$.

The above diffusion process implies that it can not influence the pixels before the tampered pixel with a grey value change. As a remedy, we here add a reverse diffusion process as a supplement to the above diffusion process.

Step 7. Iterate the following map to produce a random grey value sequence

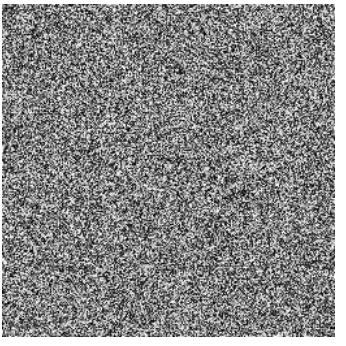
$$z_{k+1} = (z_k/c) \text{mod } 1$$
$$\phi(k+1) = \text{floor}(L \times z_{k+1}), k = 0, 1, \cdots, H \times W - 1.$$

Step 8. Execute the reverse diffusion process:

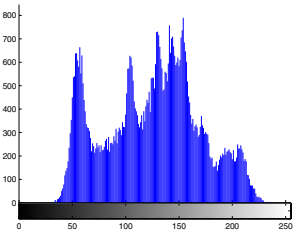
$$D(i) = D(i+1) \oplus [(C(i) + \phi(i)) \text{mod } L], i = H \times W, \cdots, 2, 1, \tag{4}$$

where $D(i), i = 1, 2, \cdots, H \times W$ are the final encrypted vector consisting of the encrypted image pixel grey-scale values. The value of $D(H \times W + 1)$ should be provided to cipher out the sequence $D(i), i = 1, 2, \cdots, H \times W$. $D(H \times W + 1)$ can handled in the same way as $C(0)$.

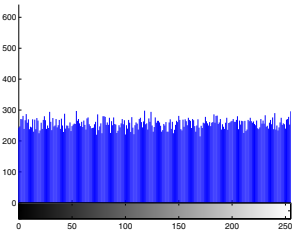
(a) The original image Lena



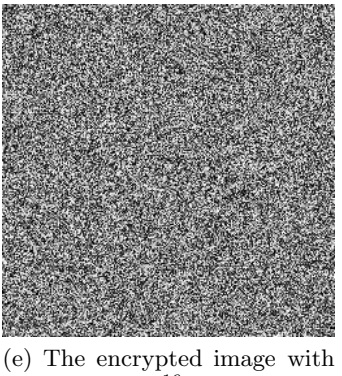
(b) The encrypted image with $a = 0.53$



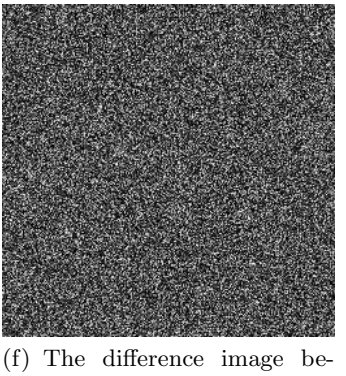
(c) Histogram of original image



(d) Histogram of encrypted image



(e) The encrypted image with $a = 0.53 + 10^{-16}$



(f) The difference image between (b) and (e)

Fig. 2. The encrypted results with $x_0 = 0.2709, y_0 = 0.7627, z_0 = 0.3607, b = 0.3216, c = 0.3902$.

The complete diffusion process is composed of Step 1 to step 8. The permutation process and the diffusion process form the proposed image encryption scheme. The image Lena is encrypted and the result is shown in Fig. 2(b).

4 Security Analysis

4.1 Key Space Analysis

Since the permutation process is irrelevant to the diffusion process, the key space consists of the cipher keys in both processes. In the permutation process, the control parameter a, the initial condition x_0 form the cipher keys. The cipher keys in the diffusion process are b, c, y_0, z_0. According to the IEEE floating-point standard, the computational precision of the 64-bit double precision numbers is 2^{-52}. Therefore the total number of different values which can be used as a is 2^{52}, so are the numbers for b, c, x_0, y_0, z_0. The key space is larger than $(2^{52})^6 = 2^{312}$. Such a large key space can efficiently prevent opponent's brute-force attack. The key sensitivity tests can be performed. In Fig. 2(e), we replace $a = 0.53$ by a perturbation $a = 0.53 + 10^{-16}$ and then encrypt the original image Lena, the yielded cipher-image has 99.59% of difference from the cipher-image Fig. 2(b) with $a = 0.53$. The difference image is shown in Fig. 2(f).

4.2 Statistical Analysis

Passing the statistical analysis on cipher image is of crucial importance for a cryptosystem. Indeed, an ideal cryptosystem should be robust against any statistical attack. In order to prove the security of the proposed encryption scheme, the following statistical tests are performed.

(i) Histogram: Encrypt the image Lena with one round, and then plot the histograms of the plain-image and cipher-image as shown in Figs. 2(c)-(d), respectively. Fig. 2(d) shows that the histogram of the cipher-image is fairly uniform and significantly different from the histogram of the original image and hence it does not provide any useful information for the opponents to perform any statistical analysis attack on the encrypted image.

(ii) Correlation of adjacent pixels: To test the correlation between two adjacent pixels, the following performances are carried out. First, we select 6000 pairs of two horizontally adjacent pixels randomly from an image and then calculate the correlation coefficient of the selected pairs using the following formulae:

$$Cr = \frac{cov(x,y)}{\sqrt{D(x)}\sqrt{D(y)}}$$

$$cov(x,y) = \frac{1}{T}\sum_{i=1}^{T}(x_i - E(x))(y_i - E(y)),$$

$$E(x) = \frac{1}{T}\sum_{i=1}^{T}x_i, \quad D(x) = \frac{1}{T}\sum_{i=1}^{T}(x_i - E(x))^2,$$

where x, y are the grey-scale values of two adjacent pixels in the image and T is the total pairs of pixels randomly selected from the image. The correlations of two adjacent pixels in the plain-image and in the cipher-image are shown in the Table 1.

Table 1. Correlation coefficients of two adjacent pixels in two images

	Plain image	Encrypted image
Horizontal	0.9435	-0.0084
Vertical	0.9680	0.0150
Diagonal	0.9157	0.0112

Table 2. Results of NPCR and UACI tests of Lena

Position	(101,200)	(164,179)	(186,7)	(179,49)	(232,180)
NPCR(%)	99.8642	99.8291	99.9252	99.9130	99.9542
UACI(%)	44.2359	42.4317	16.2499	26.9253	47.7179
Position	(242,25)	(163,255)	(39,102)	(122,24)	(248,16)
NPCR(%)	99.7910	99.8520	99.9893	99.8077	99.9390
UACI(%)	40.5671	20.8571	25.2148	41.7489	47.5361

4.3 Differential Attack

In general, attacker may make a slight change (e.g., modify only one pixel) of the plain-image to find out some meaningful relationships between the plain-image and the cipher-image. If one minor change in the plain-image will cause a significant change in the cipher-image, then the encryption scheme will resist differential attack efficiently. To test the influence of only one-pixel change in the plain-image over the whole cipher-image, two common measures are used: number of pixels change rate (NPCR) and unified average changing intensity (UACI). They are defined as

$$\text{NPCR} = \frac{\sum_{i,j} D(i,j)}{W \times H} \times 100\%,$$

$$\text{UACI} = \frac{1}{W \times H}[\sum_{i,j} \frac{|C_1(i,j) - C_2(i,j)|}{255}] \times 100\%$$

where C_1, C_2 are two cipher-images corresponding to two plain-images with only one pixel difference, W and H are the width and height of the processed image, D is a bipolar array with the same size as image C_1. $D(i,j)$ is determined as: if $C_1(i,j) = C_2(i,j)$, then $D(i,j) = 0$, otherwise $D(i,j) = 1$.

NPCR measures the percentage of different pixels numbers between two cipher-images whose plain-images only have one-pixel difference. UACI measures the average intensity of differences between two cipher-images. To resist difference attacks, the values of NPCR and UACI should be large enough. The test of the plain-image is Lena. We randomly select 10 pixels and change the grey values with a difference of 1, The numerical results are shown in Table 2. The mean values of the ten NPCR and UACI values are 99.8865% and 35.3485% respectively. It is an exceptionally good result with only one round of encryption.

5 Conclusions

An efficient image encryption scheme based on the generalized Bernoulli shift map is proposed in the paper. The proposed scheme can shuffle the plain-image efficiently in the confusion process. An effective diffusion process is also presented to change the grey values of the whole image pixels. The sensitivity to the initial conditions and control parameters and the key space are analyzed numerically. The statistical attack analysis and differential attack analysis have been performed as well. All the experimental results show that the proposed encryption system is highly secure thanks to its large key space, its high sensitivity to the cipher keys and plain-images.

Acknowledgments. This research is supported by National Natural Science Foundation of China (No. 11071152) and Natural Science Foundation of Guangdong Province (No. 10151503101000023).

References

1. Baptista, M.S.: Cryptography with chaos. Physics Letter A 240, 50–54 (1998)
2. Fridrich, J.: Symmetric ciphers based on two-dimensional chaotic maps. International Journal of Bifurcation and Chaos 8, 1259–1284 (1998)
3. Chen, G.R., Mao, Y.B., Chui, C.K.: A symmetric image encryption scheme based on 3D chaotic cat maps. Chaos, Solitons & Fractals 21, 749–761 (2004)
4. Mao, Y.B., Chen, G., Lian, S.G.: A novel fast image encryption scheme based on the 3D chaotic Baker map. International Journal of Bifurcation and Chaos 14, 3613–3624 (2004)
5. Alvarez, G., Li, S.: Breaking an encryption scheme based on chaotic baker map. Physics Letters A 352, 78–82 (2006)
6. Liu, J.M., Qu, Q.: Cryptanalysis of a substitution-diffusion based on cipher using chaotic standard and logistic map. In: Third International Symposium on Information Processing, pp. 67–69 (2010)
7. Zhang, G.J., Liu, Q.: A novel image encryption method based on total shuffling scheme. Opt. Commun. (2011), doi:10.1016/j.optcom.2011.02.039

A Design and Implementation of Simultaneous Remote Text File Editor in LAN

David Riggleman, Jeff Oakley, and Nan Jiang*

Elmer W. Engstrom Department of Engineering and Computer Science,
Cedarville University, Cedarville, OH, U.S.A.
{dkriggleman,jgoakley,njiang}@cedarville.edu

Abstract. In this paper, we present the design and implementation of a text editor that will allow for content to be remotely edited by multiple users at the same time in a Local Area Network. It is a simple, yet flexible real-time text editor that serves as a tool to help facilitate better group collaboration whether working on a programming project in plain text or a formatted report using styled text. The design is then implemented on a campus Local Area Network, which shows that this program is a very helpful tool in improving group collaboration through the minimizing of wasted time when a group member is simply waiting to modify a section of the document that no one else is currently editing.

Keywords: computer network, client/server, file synchronization, graphical user interface, local area network.

1 Introduction

A collaborative editor is a form of collaborative software application that allows several people to edit a computer file using different computers. There are two types of collaborative editing: real-time and non-real-time. Real-time collaborative editing (RTCE) is synchronous (simultaneous), meaning that users can edit the same file at the same time. Non-real-time collaborative editing is asynchronous, which means that editors do not edit the same file at the same time (similar to revision control systems).

One of the most popular and well known real time collaborative text editors is Google Docs [5]. Although it is very well designed and good for the common users, there is a limitation while working on a group programming project: all indenting is lost when trying to copy it back to a local text file after the user has finished working on the file. And since formatting is very important when working on a programming project, Google Docs would not simply work. There are a few other real-time collaborative editors that market themselves to programmers and do preserve white space. One such example is Collabedit [2]. While the syntax highlighting is a nice feature and it is a good tool for programming projects, it also has its limitations. Unlike Google Docs, which is primarily for writing formatted text, Collabedit is the opposite in that it doesn't support any text formatting.

* Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 40–47, 2011.

In this paper, we design and implement a real-time collaborative text editor that will allow for content to be remotely edited by multiple users at the same time. What our program offers is a combination of the benefits of both programs through enhanced flexibility in allowing for text to be opened and saved either as plain text or formatted text. It is simple and easy to use without the complexities found in the programs that currently exist. In addition, specifically for Cedarville students who want to collaboratively edit a file in real-time, the performance will also be better because messages sent between our server and client application will only have to traverse the Cedarville network instead of the entire Internet if going through an online real-time editors' servers such as Google Docs or Collabedit.

The paper is organized as follows: in section 1, we give an overall introduction of this paper. In section 2, we present the design of the proposed simultaneous remote text file editor. In section 3, we present the implementation of this real-time text editor communication tool. Finally, in section 4 we conclude the paper and discuss the future work.

2 System Design

In this section, we present the system design of the proposed simultaneous remote text file editor, named CedarSync. When implementing a real-time collaborative editor, there are many complexities that are necessary to think about. What will happen if multiple users try to edit the same text at the same time? How exactly do you ensure that the text and all attributes (styling information) stay synchronized? Keeping these ideas and more in our mind, we came up with the following approach to solve our program and to implement CedarSync. Beginning with the big picture, our design has several components that connect together to form CedarSync, the entire application. A class diagram of the application is shown below.

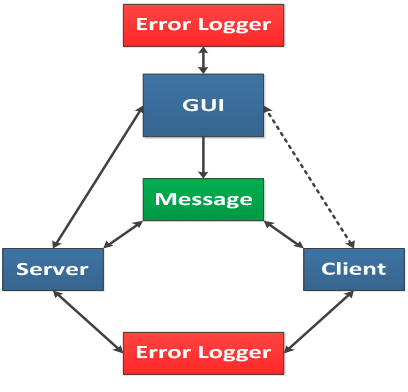


Fig. 1. Class diagram of CedarSync.

As illustrated in the figure above, first, we will have the main graphical user interface. This is the only class that the typical user will see and interface with. One important item in our design was that the same GUI could be used as either a server or a client. We did not want to have to design separate user interfaces for the two roles.

This is even more important that a user might not necessarily know at startup what role (server or client) he or she may play. This is illustrated in the class diagram as the GUI can communicate with both server and client. In the example diagram above, the GUI is currently running a server thread, indicated by the solid instead of dashed line. If the GUI was currently running a client thread, then that line would be solid and the currently solid line would be dashed. The next item to note is centrality of the Message class. This class is extremely important as it determines all communication between the server and client and even between the GUI and either the server or client threads. Due to the custom needs of an application, we defined a custom message protocol to handle synchronization messages as well as other messages such as open and close messages or chat messages. Fig.2. outlines the format of the Message class.

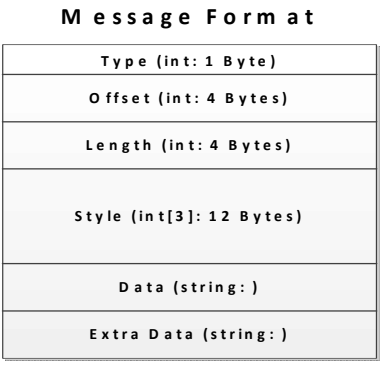


Fig. 2. Message class format

Each message object consists of several fields. The first field consists of the Message Type. The following are the valid message types: OPEN_NOTIFY (used to initialize client and server connection), REFRESH (used to indicate the server will resend all current text), INSERT (insert text), REPLACE (replace text), REMOVE (remove text), STYLE (style change), CHAT (send or receive chat message), CLOSE_NOTIFY (let connected clients know one client has closed its connection), and CLOSE_COMMAND (a command force the client to terminate its connection). The second field is the offset and it is used to indicate the position in the text file where the change was made. The length indicates the number of characters changed (inserted, replaced, removed). The style is used to set the formatting of the text (color, size, font, bold, italics, underline). The data field is used to send the actual text and therefore the size of it field can vary. The data field can also be used in non text related messages such as the open notify message to send the connected client's screen name. The final field is to allow an additional string to be sent if needed. The size is again variable and is null if not used. One example where this field is used is in the chat message, where the chat message itself is placed in the data field and the screen name of who sent the chat message is placed in the extra data field. The final class that has yet to be mentioned so far is the error logging class. This class essentially silently records all non-fatal errors that occur during program execution and writes the error message to a log file, including a timestamp of when the error

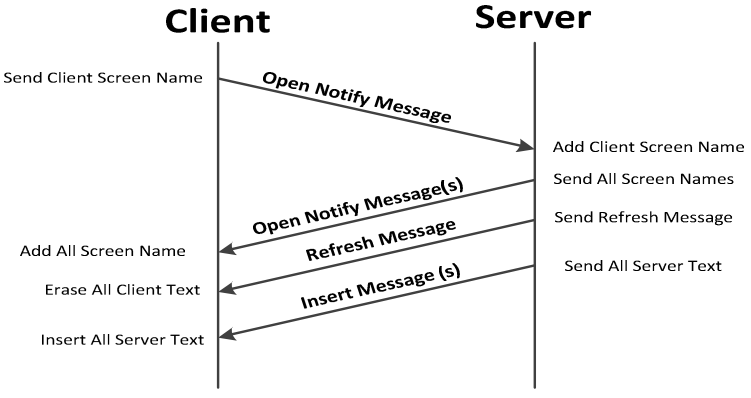


Fig. 3. Client/server connection setup

occurred. This class is very helpful when debugging or if one had to provide support to an application and the customer was experience weird problems.

The final stage of our design involved determining how the setup messages and teardown messages would work. Fig.3. shows the process that goes on when a client connects to a server. Upon connecting, a client will send an OPEN_NOTIFY message with the client's chosen screen name. Upon receiving that message, the server will add the new client to its list of connected users in the GUI. It will then send OPEN_NOTIFY messages back to that client with the server's screen name as well as all other connected clients screen names. At the same time, although not shown, the server will send a broadcast message to all currently connected clients with the screen name of the newly connected client. After received the OPEN_NOTIFY message, the server will send a REFRESH message to the client. When the message is received, the client will erase all current text in the text pane in expectation of receiving the updated text from the server. On the server side, after sending the REFRESH message, the server will send as many INSERT messages as needed (only one is needed if no styling is needed). At this point, the client's text will be synchronized with that of the server.

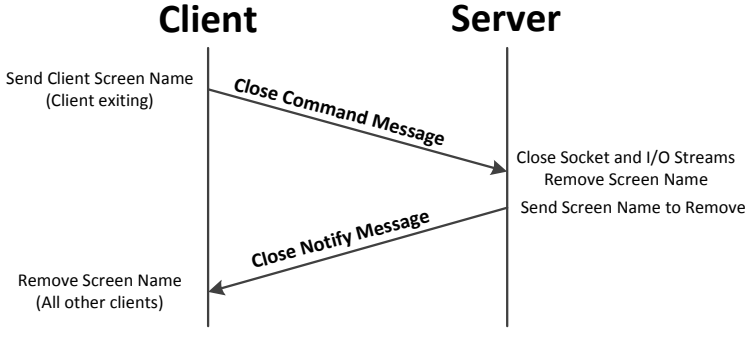


Fig. 4. Client Server connection teardown

For the teardown message, it is much simpler. As shown in Fig.4., when a client is about to exit, a CLOSE_COMMAND is sent to the server. The server, upon receiving the message, will close the client's sockets and input/output streams and remove that client's screen name from the list of connected clients. At this point, the server will then broadcast out a CLOSE_NOTIFY message to all other clients to let them know that a client has closed its connection and therefore to remove that user from their list of connected users in the graphical user interface.

3 Implementation

In this section, we present the implementation of CedarSync in the following categories: the GUI interface, message, client and server, and error logger.

3.1 GUI Interface

The GUI interface is the primarily interaction any user will have with the program. A screenshot of the user interface is shown in Fig.5.

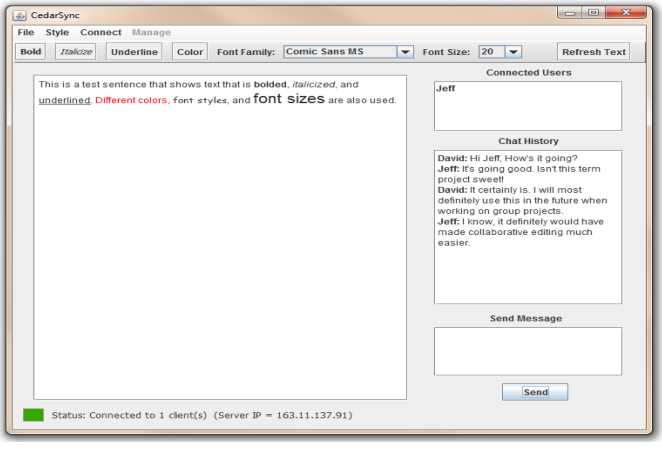


Fig. 5. GUI Interface

The main area of the GUI is obviously for editing text, since that is, after all, the main focus of the program. The text editing area is implemented in Java using a JTextPane [1]. On the right side of the GUI, is additional functionality. Near the top is a list of all users (by screen name) currently connected to the server. Below that is the area to send and display chat messages. The chat history area is a text pane in order to allow for styling (bolding the screen names). The send message area is just a text area, however, as it that is sent is the text and you cannot have styled chat messages. Moving onto the next part of the GUI, the toolbar provides convenient access to formatting and styling actions, including bold, italicize, underline, color, font family, and font size. There is also a refresh button. If currently the server, the refresh button will send a refresh message to all connected clients. If currently a client, then the

refresh button will prompt the server to send a refresh message with the server's text to just that client. The next component of the GUI is the file menu. It includes options to open and save a file in either plain text form or as formatted text. Opening and saving files in formatting texts is done through the use of an RTFKit that creates and opens .rtf (rich text format) documents that can be opened and edited in Microsoft Word or other word processing programs [1]. Opening and saving files in plain text is simply done using Java's BufferedReader and BufferedWriter classes. The final aspect of the graphical user interface is the status label at the bottom. It indicates the current status of the connection. If a client thread, the possible statuses are "Not Connected" (red) or "Connected" (green), where a connected state will display the host name (or IP address if not available) of the server. If a server thread, the possible statuses are "Not Connected" (red), "Waiting for connection on port ####..." (yellow), or "Connected to # client(s)" (green).

There are several interesting implementation details relating to the graphical user interface. The first was the need to become familiar with the SwingUtilities.invokeLater() method. In Java, there is only one main thread, the event dispatch thread, which is responsible for updating the graphical user interface [4]. However, most of the updating of the graphical user interface occurs from the server or client thread since this is where the entire message sending and receiving is happening. As a result, there could be more than one thread modifying the GUI, which could potentially lead to deadlock. As a result, whenever there is an update to the GUI that could occur from an outside thread, the code needs to be run using the invokeLater method, which adds the GUI update event to the event dispatch thread's queue.

Second, the messages are sent from the graphical user interface to the server or client which then sends the message. This works through the use of a Document Listener class within the GUI class. Whenever a change to this document is made, an event triggers with the given information that will be encoded in the message, such as the offset position, the length of the change, and the new text. Because the GUI cannot directly communicate with the server or client since they are implemented as threads, the GUI adds the message to a shared queue with the client and server thread and then notifies that thread that there is a new message ready to be sent. One other thing that we discovered was that it is necessary to disable the document listeners when modifying the text based on a received message. If the listeners are not disabled, then the received event will be treated as a new event, and the servers and clients will infinitely send the same message back and forth until the text pane overfills with text and the maximum number of characters allowed is reached.

3.2 Message

Since an extended discussion of the message class occurred in the design section, there is no need to repeat it here as the implementation was according to the design. Instead, we will discuss a class contained within the Message class file, the StyleConverter class, that is used to convert the style into a format that be sent across the network. As previously mentioned, Java supports serialization for most objects [6]. Unfortunately, attributes are not one of those classes and cannot be serialized. As a result, the StyleConverter class contains two methods, one that encodes an attribute into an integer array and another that decodes the integer array to an attribute. The integer array to store the attribute consists of three integers. The first integer is used to

store three style items: bold, italics, and underline. This is done through the use of setting and reading specific bits to indicate different values; the first bit sets the bold attribute, the second bit sets the italic attribute, and the third bit sets the underline attribute. The second integer is then used for the font color, converted to an integer. The final integer is then used to store the font size. If left wondering what about the font family attributes, this is actually sent in the extra data field as a string.

3.3 Server/Client

As there are some similarities in the Server and Client classes, we will first discuss the implementation details that are common to both before talking about the details unique to each class.

One of the challenges in implementing CedarSync was the communication that had to occur between the main GUI class and the server or client thread. As aforementioned, the solution was to use a shared linked queue. Then whenever the GUI had a message to send based on a document change or a new chat message, a new message was added to the queue. The challenge in the threads then is in knowing when there are new messages ready to send. As a result, it was necessary for both the client and server threads to spawn a new thread whose sole task it is to listen for messages from the GUI and then pass them onto the respective client or server send function. Whenever the GUI added a new message to the thread it would notify the thread object that there was a new message to send. This would cause the thread object in the client or server thread to wake up, at which point it would process all messages that were currently in the queue until the queue was empty. That message listening thread would then call the wait() method, which would put the thread to sleep until the notify method is called again by the main GUI when there is another new message to send.

Moving onto the specifics of the Client class, besides the additional message listening thread mentioned above, the Client class is not much different than the typical TCP client [2]. The only real difference then is in the process messaging function. In most cases, upon receiving a message, the client will extract the relevant information and call the appropriate GUI method, such as updating the text pane or adding a message to the chat history pane.

Finally, with the Server class, there are several more implementation details. The first occurs in the beginning when the server is first started. The constructor to the server class takes a port number, which is the port number the user selected from the dialog on the main GUI. If, however, the port number is currently in use, the program will automatically and randomly select a new port number to use, attempting up to a specified number of random numbers, currently set at 10. Another important item to note about the server class is the data structures that are used to store client information. Whenever a new client is accepted, the both the clients socket and the client's sockets are added to each ArrayList. When a client then terminates its connection, the client's socket is removed from the socket ArrayList while the client's output stream is removed from the output stream ArrayList. Next, there is a HashMap where the key is a screen name and the value is the socket related to that screen name. The HashMap is used for sending messages for a specific client, as to send a message to a specific screen name simply involves getting the client's socket by plugging the screen name into the HashMap.

3.4 Error Logger

The error logging class is simple, but very useful and can be used with all other classes, thus only needing to be implemented once. The constructor for the ErrorLogger class takes the desired filename of the error log. If no name is specified, the default error log name is "errorLog.log" The class then contains one method called writeToFile that takes a string with the message wanting to be written to the error log. The method will first write to the file the current date and time before writing the error message passed in as a parameter. Finally, there is a close method that simply closes the log file.

4 Conclusion and Future Work

CedarSync has performed according to our proposal, and experience from on campus network implementation shows that this program will be a very helpful tool in improving group collaboration through the minimizing of wasted time when a group member is simply waiting to modify a section of the document that no one else is currently editing. Our program also has a very low learning curve as it is very simple to use. In the future, we plan to add some additional features to this program. The first would be making the connection secure by encrypting the message (using Java's Cryptography classes) and also forcing a connecting client to authenticate himself with the server. The other possible feature would be to lock specific parts of the file from other users editing that section, but we ultimately felt that the time needed for this feature would not be worth the questionable benefits, as well as the added complexity that we were trying to avoid as previously mentioned. We hope and believe many people will benefit from the enhanced collaboration that CedarSync provides.

References

1. Java JTextPane Save. Java Programming Forums (2010), http://www.javaprogrammingforums.com/file-i-o-other-i-o-streams/4218-java-jtextpane-save.html#post11565 (retrieved December 5, 2010)
2. Nolan, B.: Online Collaborative Text Editor. collabedit (2008), http://collabedit.com/ (retrieved December 5, 2010)
3. Peterson, L.L., Davie, B.S.: Computer networks: a systems approach, 4th edn. Morgan Kaufmann, Amsterdam (2007)
4. Threading with Swing: SwingUtilities.invokeLater. Javamex (2010), http://www.javamex.com/tutorial (retrieved December 5, 2010)
5. What's new in Google Docs. Google Docs (2010), http://www.google.com/google-d-s/whatsnew.html (retrieved December 5, 2010)
6. Zukowski, J.: Advanced Serialization. Sun Developer Network (2001), http://java.sun.com/developer/technicalArticles/ALT/serialization/ (retrieved December 5, 2001)

The Storage of Weighted Directed Graph in Relational Database and Implementation of Dijkstra Algorithm

Yanchen Li[1,2],Yuan Feng[1,*], Feng Hong[1], and Ying Bi[3]

[1] College of Info. Sci. & Tech.Ocean University of China, Qingdao
[2] Hisense Trans-Tech Co,Ltd Qingdao, China
[3] College of Business Administration, QingDao Hismile College

Abstract. As an important data structure, map is widely used in modern enterprise management software. Route planning, decision support, finding the critical path, and other activities all need the algorithms in graph theory. And once the map becomes more complex, it has to be stored in database. While the traditional adjacency matrix and adjacency list stored in the database have their limitations. If still achieve the appropriate algorithm using a programming language, software and database will exchange large amounts of data when the software and database are not in the same computer. So this method will inevitably affect the efficiency of the algorithm, and this situation is extremely difficult to optimize. Based on the characteristics of relational database, our paper designs a storage model of a directed graph in relational database. Using the SQL language and set operations, we successfully achieve the Dijkstra algorithm. In the process of Dijkstra algorithm, we use the mechanism of UPDATE in SQL instead of the inner loop steps in Dijkstra algorithm, thus the time cost by database operations is reduced greatly and the efficiency is raised.

Keywords: Weighted directed graph, Relational database, Dijkstra algorithm, Sql.

1 Introduction

As an important data structure, map is widely used in modern enterprise management software. Route planning, decision support, finding the critical path, and other activities all need the algorithms in graph theory. Once the map becomes more complex, it has to be stored in database. The traditional adjacency matrix and adjacency list stored in the database have their limitations:

1. For the storage of adjacency matrix, since the number of nodes and edges are different, so the numbers of rows and columns representing the adjacency matrix are different. In such situation, you need to store all the tables as many as all the maps, which is not conducive to database performance optimization. And this method makes the comparison between maps different. When the weights of the edges become more complex, it is harder to describe the compositions of the weights of edges. We can only record the results calculated as the weight, as a result, analysis the weights calculated dynamically based on the user' demand becomes difficult.

* Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 48–54, 2011.

2. If you use adjacency list, its essential pointers in the relational database is difficult to achieve, common method to solve this problem is to take a whole row as a pointer and store the ID of the corresponding line, according to the stored ID we can find the corresponding row. But such method needs to use Recursive queries and cursors. And in relational database, its implementation is inefficient.

In the two storage modes above, Dijkstra used in map analysis is usually achieved by programming language, which will cause large amount of exchange between applications and databases. Once the software and database is not in the same computer, the network speed may become operational Bottleneck, then the system will be very difficult to optimize.

Dijkstra using relational database often takes database language as programming language, it uses application variables and map cursor to operate map step by step, so its query is not efficient. If the transaction is busy, it will easily lead to deadlock.

According to the characteristics of a relational database, this paper designs a storage model using only one table to store number of directed graphs with same nodes. This model can store a variety of weights without pointers. The implementation method is easy to use database to implant the horizontal comparison between maps and facilitate the dynamic calculation of weights. Most graph operations can be achieved through a database set operations, which can greatly reduce the interaction between programming language and databases and improve the efficiency of the system.

1. On the base of such storage, we use SQL language to achieve Dijkstra algorithm. Using the mechanism of UPDATE in SQL language instead of the Dijkstra algorithm inner loop, we can find and record the whole diagram of the vertices of the shortest paths with only one loop.

2. Replace the use of the programming languages, using database language to achieve Dijkstra can reduce the large times of interactions between the database and program, which improves the efficiency and reliability greatly.

2 Table Structure and Algorithm Design

2.1 The Storage of Directed Graph in the Database

In the process of storing graph using adjacency list, we find that some particular collection of database operations can directly find the desired data. Start point and end point can be recorded to represent an edge, and single node can be seen as an edge from point to their own. In order to distinguish arcs their own and nodes, we can use a flag. The arc right and the field of DATA can be expressed using the same column.

A weighted directed graph can be expressed in Table 1(Table 1 is defined as Graph, in order to facilitate the use of SQL statements).

Table 1. Graph

BgnNode	EndNode	Type	Data
V0	V0	NODE	0
V1	V1	NODE	1
V0	V5	SIDE	100
……	……	……	……

In Table 1, BgnNode column represents the starting of edge, EndNode column represents the end of Edge. Type column is marked to indentify the node and edge. If the Type is node, the data of BgnNode and EndNode must be the same. Data column stores the weights of edges or the data information of nodes.

Using the method introduced in this section to store map, operations become easier, which can use SQL set to achieve.

As the actual application has no need of traversal algorithm, so this paper does not do research on traverse algorithm, but we realize the Dijkstra algorithm.

2.2 The Operation to Achieve Djistra Using SQL

Dijkstra algorithm is useful in many cases, the storage method described in this paper can use the SQL set operations to support the implementation of the Dijkstra algorithm.

Dijkstra is an algorithm that finds the shortest path from one point to other points. The basic idea of the algorithm is:

Let V to be the set of all vertices in the graph, the source v0 belongs to the set V;

Let set S to be the set of vertices having found the shortest path, Suppose T = V-S.

Select the vertices of shortest path to v0 from the collection T, and add the point u to set S. Every time adding point u, we modify the value of the shortest path from v0 to other points in set T. The length of other vertices in T is the sum of the original value and the shortest path values of u and the smallest path value of this point. Repeat this process until all vertices of T have added to S[1].

2.2.1 The Design of Result Table

In order to implant Dijkstra in the database, we need a table to store the results of the algorithm execution.

It can be seen from the algorithm, the algorithm results of each cycle will affect the next cycle, so there should be a flag in the table to identify whether a particular node has found the shortest path. We make the flag of the node has found the shortest path to be 2, and otherwise we take the flag to be 0.

Table should have a column to store nodes on the shortest path and the total length of the whole path.

According to the representation of maps, there should be the starting point and end point in the table.

Based on the Dijkstra, we can let the length of vertexes yet to find the shortest path to be the maximum that the computer allowed. In this paper, we use the maximum 65534.

A structure of a result table is shown as table 2.

Table 2. Djistra

BgnNode	Path	EndNode	PathLength	Flag
V0		V0	65534	0
V0		V1	65534	0
V0		V2	65534	0
……		……	……	……

BgnNode is the starting point and EndNode is all the nodes in the table. So all the paths described in the paper are recorded. The path can be constituted by a sequence of vertices, which is recorded in the feild of Path. PathLength field is used to store the length of the path.when the shortest path has not exited, the value is recorded to be maximum 65534 and the FLAG is 0. When the Flag is 1, it presents that the state is the middle state of finding the shortest path.

2.2.2 Algorithm Design

Seen from the description of the Dijstra, the algorithm is composed of two loops, inner loop will find the current shortest path to all the vertices. Every time finishing the outside loop, it will find the shortest path to add into the collection of S. Next time, the vertexes of set S will be the initial conditions, and then start the inner loop.

The mechanism of UPDATE in SQL is to find the appropriate conditions after scanning the whole table, and then update the relevant records. Properly set the conditions, it can certainly replace the inner loop. Besides, you can optimize with the use of database indexes, whose performance is superior to the circulation structure compiled by its own.

Outline design of the algorithm is as follows:

1. Based on the Dijkstra, after initialized the relevant variables, the first step is to check whether there are still nodes that have not found the shortest path. If not, then exit the program directly; otherwise, start the loop.
2. The next step, it should find the shortest path from the source to the whole nodes can be reached, and add it into result table. This is the function of inner loop. In this paper, we use a UPDATE statement to achieve.
3. After getting all paths, we choose the shortest path not belonging to S; add it to set S and set the flag to be middle state. Such node will be added to set S.
4. Next, we store the variables of intermediate state nodes, which will be used in the next loop.
5. Set the intermediate state of the vertexes to 2, and then end this loop to start the next loop.

Stepwise refinement:

It can be seen from the outline design and flow chart, the algorithm focuses on how to use the UPDATE statement instead of inner loop of the Dijkstra algorithm.

Update statement can be divided into two aspects: the searching conditions and the content of records.

Suppose there is an edge from V_i to V_j, we set its weight to be PathLength (V_i,V_j), the source node is V_0.

1.Searching conditions:

After the first n cycles, we defined the vertice added to set S to be vi. Then the next loop, we should find all the vertices V_x to V_i. And PathLength (V_0,V_i)+PathLength (V_i,V_x) < PathLength (V_0,V_x).

2.content of records

First we should record the weight of the shortest path, not only for the final result, but also for the process of the algorithm step by step. Principles of records are:

Set the path length of point V_i which is the last added to set S to be ShortestPathLength, if ShortestPathLength+PathLength (V_i,V_j)<PathLength (V_0,V_j) exited, then we record the path length of V_j as ShortestPathLength + PathLength (V_i,V_j).

Second, the process should record the vertices on the shortest path:

Defined the node added to S to be Vi after n-1 loops, the node of the next loop is Vj, the in the (n+1)th loop:

If Vj=V0, then there is no need to record the path. Such situation is there is an edge between V0 and Vx, and this edge is the shortest path.

If there is an edge between Vi and Vj and we find the edge between Vj and Vx in this loop, then the shortest path to Vx is the extension of the path from V0 to Vj.

Otherwise, Vj is the first node on the shortest path from V0 to Vx.

Seen from the factor above, which parameters the UPDATE statement requires is:

1. The first one is length of the shortest path of current vertexes, which are recorded in the result table.
2. The second one is the shortest path of the node added to S last loop.
3. The next one is the flag of the point Vj added to S in the last loop.
4. The forth one is the flag of the point Vi added to S in the loop before the last loop.
5. The last one is to determine whether there is an edge between Vi and Vj.

Such parameters should be stored into variables, which is for the next loop.

3 Structure of the Table and Implementation of the Algorithm

3.1 The Storage of the Graph

According to the design of Section 2.1,the SQL that the table used to store the Graph is as followed:

```
CREATE TABLE  graph (Bgnnode  varchar (10) not null,
  endnode varchar (10) not null, type varchar (50) ,
                                 data  numeric(18, 0) )
```

If the two nodes have the same direction edge with different weights, we can take the data field as primary key, using weights to distinguish the edges.

If the weights are constituted more complexly, we can extend the data field to multi-column.

If you want to store the graph with same nodes and different edges and different weights, we can add a field to uniquely identify a graph, such as GraphCode, and then add it to primary key.

After storing the graph, the next operations are described in section 2.2.

3.2 The Algorithm of Dijkstra

3.2.1 The Design of the Result Table

According to section 2.3.1, the SQL statements of result table is as shown:

```
CREATE TABLE dijkstra ( Bgnnode varchar(10),
Path varchar (200) null,Endnode varchar(10) ,
PathLength numeric(18, 0),flag varchar (1) not null)
```

Field Path will store the nodes on the shortest path. In this paper, we do not insert separator since the length of the node ID is the same. If needed, we can modify the algorithm UPDATE.

If you execute few graphs with same nodes, Dijkstra can add a unique identifier field into the table.

3.2.2 Implementation of Dijkstra

Prerequisite for the implementation of the algorithm is to enter a starting node, which is often used as storing procedure parameters in practice. This paper defines a variable "@startnode" instead.

```
/* initialize the result table */
Insert into dijkstra
select @startnode as bgnnode,endnode,'0' as flag,'' as
path,65534 as pathlength from Graph where Type='Node'
/* initialize the variable */
declare @vs_bgn varchar(5) set @vs_bgn=@startnode
declare @vi_PathLength numeric(18,0)
    set @vi_PathLength =0
declare @vs_last varchar(5) set @vs_last=@startnode
/* Next start writing the robin algorithm */
while (select distinct 1 from dijkstra where flag=0)=1
/* The contents of the loops are described below
   The first step: Update the results table */
update dijkstra
set PathLength =b.data+@vi_PathLength,
    Path=(case when @vs_bgn=@startnode then ''
               when @vs_bgn=@vs_last then a.Path+@vs_bgn
               else @vs_bgn  end)
from dijkstra a,graph b
where a.endnode=b.endnode and b.type='side'
and b.bgnnode=@vs_bgn
and b.data+@vi_PathLength <a.PathLength
/* The second step, find the shortest path */
update dijkstra set flag='1'
where PathLength =(select min(PathLength) from dijkstra
where flag='0')
```

```
/*the third step, store the relevant parameters into variables */
select @vs_last=b.endnode from dijkstra a,graph b
where a.flag=1 and b.bgnnode=@vs_bgn and
b.endnode=a.endnode and b.type='side'
select @vs_bgn = endnode  from dijkstra where flag=1
select @vi_PathLength = PathLength from dijkstra where
flag=1
```

```
/*The fourth step, set the middle state to be 2, presenting the state have found
the shortest path.*/
update dijkstra set flag='2' where flag='1'
```

The End

4 Summary and Outlook

The storage structure in our paper is still a structure of chain, which is a combination of adjacency list and reverse adjacency list. Giving full consideration to the characteristics of the relational database, we use the set of the SQL operations to implement the map algorithms. This storage method makes the operations more simple and easy to understand. Using database greatly reduces the interactions between software and database, which avoids the bottleneck of the network and raises the efficiency.

Adding a coding field, a table can store many different types of maps. Such method has good compatibility, and it is easy to take comparison between graphs. Besides, modifying the relevant operations can achieve bulk operations of graphs efficiently. And, such storage can also store the structure of tree and list.

Dijkstra and the storage proposed by this paper have played an important role in the research of the distribution system of a global procurement center, whose performance is greater than the traditional ways.

In order to ensure backward compatibility, this database environment is the MS SQL2000, whose syntax is slightly different from under other databases under. Whether such method is support the minimum spanning tree, depth first search, finding the longest path and finding critical path and so on still need further research.

Acknowledgement

This research is supported by Research and Teach Program of Ocean University of China under grant number of 2009JY10.

Reference

1. Deng, W., Li, Y.: Data Structure (C-language version), 2nd edn., pp. 108–111. Electronic Industry Press, Beijing (2007)

Video Broadcasting over Heterogeneous Networks

Zhuo Gao[1,⋆] and Xuemei Sun[2]

[1] Beijing Information Science & Technology University, Beijing 100192, China
gaodigang@bistu.edu.cn
[2] California State Polytechnic University, Pomona, CA 91768, USA
xuemeisun@csupomona.edu

Abstract. A new scheme is proposed for broadcasting video signals over heterogeneous networks. The scheme is based on multiple-description coding (MDC) and aims to suppress the threshold effect. Each description is independently generated and can be used alone or together with other descriptions to reconstruct the video signal. The performance tradeoff is analyzed in an information theoretic manner and it is shown that the performance of the proposed scheme approaches the optimum rate-distortion bound under certain source and channel assumptions. Practical simulation results with video signals are provided and discussed.

Keywords: video broadcasting, heterogeneous network, multiple description coding, threshold effect suppression.

1 Introduction

Recently, with the rapid development of mobile devices, there is a fast-growing demand of video transmission over a heterogeneous network. Particularly, a large number of users could request the same contents almost simultaneously, which is a typical scenario of video broadcasting. For instance, millions of users around the world watched the opening ceremony of Beijing Olympics in 2008, and the users have different reception devices, such as television sets, computers, mobile phones, etc. The heterogeneity of the network lies in several aspects: First, the connection speeds of different users vary vastly. A DSL user might be able to watch the high-definition streaming seamlessly, whereas a mobile user on a cellular network could only watch a low bit-rate version. Second, different devices have significantly different computing power, which, in turn, decides whether the complex modern video codes can be decoded in real time. Third, due to the limitation of screen size, different devices require different resolutions. There are possibly other distinctions among users, but we focus ourselves on the aforementioned aspects.

⋆ The work of Zhuo Gao is supported by Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality (PHR201108268).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 55–59, 2011.

Closely related to the broadcasting problem is the (in)famous threshold effect, also called cliff effect. It describes the sudden loss of the digital signal when the channel quality deteriorates, instead of the gracefully fading of the analog signal. In video broadcasting networks, the threshold effect appears as that all the users whose channel quality is inferior to the designated quality can not decode the video signal at all, whereas all other users have the same reconstructed video, i. e., the users with better channels do not have any performance gain.

Currently, video delivery networks usually use multiple unicasts to imitate broadcast by providing every user with a customized stream. This approach poses several problems, such as the obvious waste of bandwidth and energy to transmit the same content multiple times. More importantly, even one user can have varying channel condition, especially in the wireless environment. Thus a universal broadcasting scheme is much desirable. Recognizing the limitations and drawbacks of traditional transmission schemes, several new schemes have been proposed. The H.264 standard has Scalable Video Coding (SVC) extension [1], which is a typical progressive source coding design. The video signal is encoded into different layers with decreasing importance. The most crucial base layer has to be decoded before any refinement layer can be made use of. Microsoft$^{\text{TM}}$ proposed smooth streaming [2], which requires proper estimate of channel quality and huge amount of extra storage for different bit-rate versions of the same video content. In [3], researchers from MIT proposed SoftCast, which requires extremely large bandwidth, since it is essentially uncoded transmission of transformed video signals.

In this paper, we propose a multiple-description based scheme to broadcast the video signals. In contrast to the progressive layering schemes, every layer or description is statistically equal, and thus there is no special base layer and any subset of descriptions can lead to reconstruction of the signal. As the user's channel quality improves, more descriptions will be received and the reconstruction quality improves consequently.

2 Multiple Description Coding Model

The general source signal is modeled by a binary random variable $X \in \{0, 1\}$ with the uniform probability mass function $P_X(0) = P_X(1) = \frac{1}{2}$. The channel is a binary symmetric channel (BSC) with crossover probability p, where p varies and $p \in [0, \frac{1}{2}]$. The capacity of the channel is then given by $C = 1 - H(p)$, where $H(p) = -p \log p - (1 - p) \log(1 - p)$ is the binary entropy.

The source is encoded into k different descriptions, each by an optimum source encoder characterized by the backward test BSC with crossover probability p_q. The source coding rate is decided by the largest possible channel capacity and number of descriptions k. For the sake of simplicity, we encode the source into a sum rate of $C_{\max} = 1 - H(p_{\min})$ bits (per source sample), and therefore each description has a rate of $\frac{C_{\max}}{k}$. Each description is then encoded with an optimum channel code. It is obvious that all the descriptions are independent and each of them can be used to reconstruct the original signal.

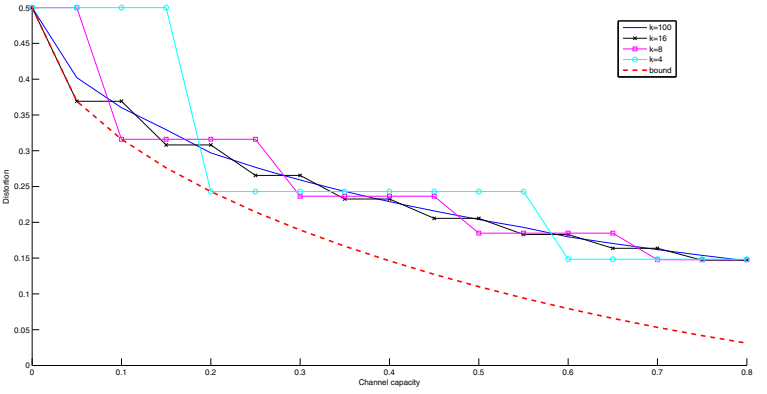


Fig. 1. Distortion performance of multiple description of binary source over BSC

Some of the descriptions could be lost during transmission as a result of channel corruption or packet loss. When m out of the k descriptions are recovered, a majority voting mechanism is employed for the reconstruction. The resultant Hamming distortion is given by the following lemma.

Lemma 1. *The Hamming distortion D as a function of the number of received descriptions m is*

$$D(m) = \begin{cases} \sum_{i=\frac{m+1}{2}}^{m} \binom{m}{i} p_q^i \bar{p}_q^{m-i} & m \text{ is odd} \\ \sum_{i=\frac{m}{2}}^{m} \binom{m}{i} p_q^i \bar{p}_q^{m-i} - \frac{1}{2}\binom{m}{\frac{m}{2}} p_q^{\frac{m}{2}} \bar{p}_q^{\frac{m}{2}} & m \text{ is even} \end{cases},$$

where $\bar{p}_q = 1 - p_q$ and $p_q = H^{-1}\left(1 - \frac{C_{\max}}{k}\right)$.

Proof. When $m = 0$, a random guess will give the trivial distortion of $D(0) = \frac{1}{2}$. When $m = 1$, notice that the description is encoded with an optimum channel code to match the rate of $\frac{C_{\max}}{k}$, which indicates

$$C_{\max} = k[1 - H(p_q)]$$

and immediately translates to $p_q = H^{-1}\left(1 - \frac{C_{\max}}{k}\right)$.

So it is obvious that when $m = 1$, the Hamming distortion $D(1) = p_q$. When m is odd, the reconstruction error happens when more than (and including) $\frac{m+1}{2}$ descriptions are flipped. The resultant distortion is then given by

$$D(m) = \sum_{i=\frac{m+1}{2}}^{m} \binom{m}{i} p_q^i \bar{p}_q^{m-i}.$$

Similarly, when m is even,

$$D(m) = \sum_{i=\frac{m}{2}}^{m} \binom{m}{i} p_q^i \bar{p}_q^{m-i} - \frac{1}{2}\binom{m}{\frac{m}{2}} p_q^{\frac{m}{2}} \bar{p}_q^{\frac{m}{2}}.$$

In Fig. 1, the distortion curves against the effective channel capacity are shown with $k = \{4, 8, 16, 100\}$. The rate-distortion bound is also plotted as a reference. As can be seen in the figure, when the effective channel capacity coincides with $\frac{C_{\max}}{k}$, the multiple description coding scheme achieves the bound, whereas the distortion performance degrades gracefully in other cases. Thus the multiple description coding scheme can adapt to varying (with one user) or different (with multiple users) channel conditions. The adjustment happens spontaneously without the help of channel feedback which is necessary for some other schemes. However, the gain of graceful degradation is obtained at the cost of performance loss at the desired channel capacity except when it is close to $\frac{C_{\max}}{k}$.

(a) Original frame, 8bpp (b) Compressed with 1 description, 1.19bpp

(c) 2 descriptions, 1 received, 0.52bpp (d) 2 descriptions, 2 received, 1.08bpp

Fig. 2. Comparison between different methods

3 A MDC Video Broadcasting Scheme

Following the same idea, we can independently generate multiple descriptions of the video signal by exploiting its temporal or spatial redundancy [4]. To demonstrate this idea, we employ a conceptually simple MDC video broadcasting scheme described as follows.

Two spatially correlated descriptions are created by sampling the pixels of a frame, and each of the description is compressed with a routine wavelet-based compression algorithm. Besides simplicity, the advantage of the scheme is that the creation of multiple descriptions is independent with both the source and channel coding components. Although wavelet-based compression is adopted here, it does not affect the relative performance when other compression algorithms are imposed. The results are shown in Fig. 2.

As we can see from Fig. 2, the source is a frame taken from the test video "foreman", which is 352×288 CIF format in grayscale. The uncompressed frame, shown in Fig. 2(a) uses 8bpp and with the same compression algorithm, the frame is compressed into 1.19bpp, as shown in Fig. 2(b). Two descriptions of the frame are created with the proposed scheme and when the channel is corrupted so that only one of the two is received, the resultant reconstruction of the frame is shown in Fig. 2(c). The video quality is low as expected, but the (approximate) half of the channel capacity is made good use via only one description and the threshold effect is suppressed. When both descriptions are received, shown in Fig. 2(d), the quality of the reconstructed video is better than the counterpart of one description received, but worse than that with *only* one description (traditional coding).

This, again, indicates that multiple description coding gains robustness at the cost of some performance, and thus when the channel condition varies dramatically (or equivalently, many users have various channel conditions), MDC based schemes are competitive choices.

4 Conclusion

We propose a novel video broadcasting scheme for heterogeneous networks. The new scheme generates multiple descriptions of the video signal and successful reception of any subset of the descriptions leads to reconstruction of the original video signal. The reconstruction quality improves with the number of the received descriptions. As the next step of our work, inspired by the results in [5], we are considering the hybrid of digital multiple-description and uncoded transmission to take advantage of both sub-schemes.

References

1. Schwarz, H., Marpe, D., Wiegand, T.: Overview of the scalable video coding extension of the H.264/AVC standard. IEEE Trans. Cir. Sys. for Video Tech. 17, 1103–1120 (2007)
2. Zambelli, A.: Smooth Streaming Technical Overview. Microsoft Corp. (2010), http://learn.iis.net/page.aspx/626/smooth-streaming-technical-overview/
3. Jakubczak, S., Katabi, D.: Softcast: one-size-fits-all wireless video. In: SIGCOMM 2010, pp. 449–450 (2010)
4. Goyal, V.: Multiple description coding: Compression meets the network. IEEE Signal Processing Mag. 18, 74–93 (2001)
5. Gao, Y., Tuncel, E.: New Hybrid Digital/Analog Schemes for Transmission of a Gaussian Source Over a Gaussian Channel. IEEE Trans. Inf. Theory 56, 6014–6019 (2010)

Studies on TBRPF Routing Algorithm and Detection of Routing Loop Attack in Mobile Ad-Hoc Network

Deboleena Ghosh

Student of Calcutta Institute of Engineering and Management,
Under West Bengal University of Technology, Kolkata, India
deboleena16@yahoo.com

Abstract. At present, the application Mobile Ad-hoc Network (MANET) is increasing at a steady pace. The basic problem with MANET lies in its ad hoc nature. The lack of infrastructures in MANET makes the detection and control of security hazards all the more difficult. In this paper, an attempt has been made to study the various attacks in MANET. The focus is mainly on the routing-loop attack which is one of the possible attacks in ad hoc networks. In a routing-loop attack, a packet may circulate around a loop for some time before its TTL (Time To Live) or Hop count field goes to zero. Even when only a small portion of the traffic enters these loops and only for a brief time, effect is impact on surrounding medium and traffic, thus degrading network performance. In order to prevent this kind of attack, it is crucial to detect the abnormality that occurs during the attack. In conventional schemes, anomaly detection is achieved by defining the normal state from static training data. However, in mobile ad hoc networks where the network topology dynamically changes, such static training method could not be used efficiently. In this paper, I propose an anomaly Detection scheme using dynamic training method in which sequence number of each packet in transit is calculated each time to discard an errant packet.

Keywords: MANET, TBRPF.

1 Introduction

A mobile ad hoc network (MANET) is a dynamic, wireless network that can be formed without any pre-existing infrastructure. MANET has no clear line of defense, so, it is accessible to both legitimate network users and malicious attackers. One of the main challenges in MANET is to design a robust security solution that can protect MANET from routing attacks. Different mechanisms have been proposed using various cryptographic techniques to counter-measure the routing attacks against MANET. However, these mechanisms are not suitable due to MANET resource constraints, i.e., limited bandwidth and battery power.

A mobile ad hoc network, sometimes called a mobile mesh network, is an infrastructure-less, self-configuring network of mobile devices connected by wire-less

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 60–66, 2011.

links. [1]Thus a MANET is an autonomous system (AS) of mobile nodes with no centralized controller. Each device in a MANET is free to move independently in any direction, and will therefore change its links to other devices frequently. Each must forward traffic un-related to its own use and, therefore, be a router. Topology of the distributed network is time varying. The dynamic nature of the network topology increases the challenges for the designing of ad hoc networks.

1.1 Attack in MANET

1.1.1 Blackhole

It has two properties. First, the node exploits the mobile ad hoc routing protocol (as AODV), to advertise itself as having a valid route to a destination node, even though the route is spurious, with the intention of intercepting the packets. Second, the attacker consumes the packets without any forwarding.

1.1.2 Resource-Consumption or Sleep-Deprivation

An attacker or a compromised node can attempt to consume battery life by requiring excessive route discovery or by forwarding unnecessary packets to the victim node. The limitedly available resources are targeted as bandwidth, computational power.

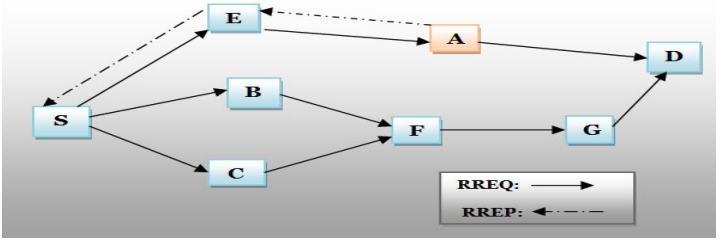


Fig. 1. Blackhole attack on AODV

1.1.3 Location Discovery

An attacker reveals information regarding the location of nodes. It gathers the node location information as a route map and then plans further attack scenarios. Traffic analysis, one of the subtlest security attacks against MANET is unsolved.

1.1.4 Wormhole

It is a network layer attack. [2]A malicious node receives packets at one location in the network and tunnels them to another location in the network where these packets are resent into the network. This tunnel between two colluding attackers is referred to as wormhole. The attacker may create a wormhole even for packets not addressed to itself because of broadcast nature of the radio channel.

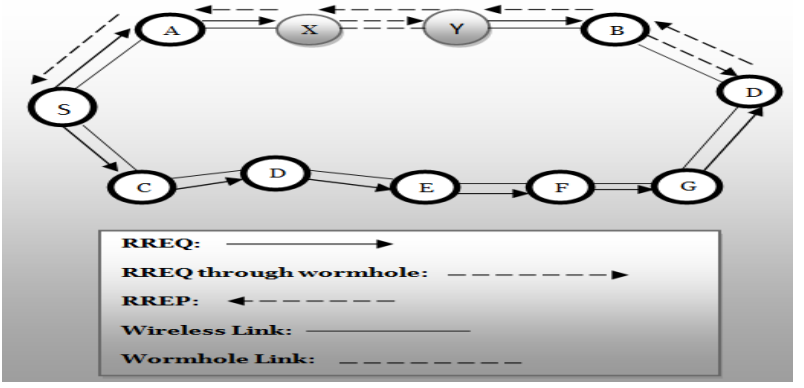


Fig. 2. Wormhole attack on AODV in MANET

1.1.5 Routing Loop Attack

To maintain loop-free routes to destination, it is required that the same topological image is perceived between neighboring nodes and thus across the network as a whole and to make routing decisions which are consistent throughout. However, transient routing loops form in ad hoc networks running pro-active routing protocols of MANETs, using hop-count metric. These loops are not the result of malevolent nodes, but those that are unable to maintain a coherent and up-to-date perception of the network topology. Even when only a small portion of the traffic enters these loops and only for a brief time, effect is impact on surrounding medium and traffic, thus degrading network performance. A packet may circulate around a loop for some time before its TTL (Time To Live) goes to zero. Due to mobile nature of mobile ad hoc networks and transient nature of the wireless medium, links form for a short time scale.

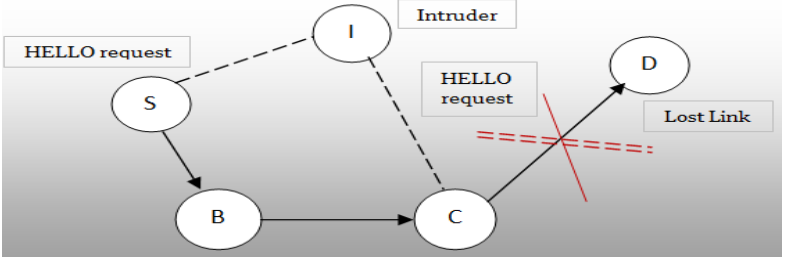


Fig. 3. Routing Loop

1.2 Routing Protocols in MANET

1.2.1 Destination-Sequenced Distance Vector (DSDV)

DSDV is a traditional table-driven protocol for MANET. Each node maintains one or more tables that contain route information to other nodes in the network. Nodes continuously update the tables to provide fresh view of the entire network. Updates

are so frequent that advertisements must be made regularly. [3]The data that is broadcast by the mobile node contains its new sequence number, destination address, number of hops needed to reach the destination and sequence number of the information received for the destination. The fundamental issue is creation and maintenance of the tables which need to be frequently updated by transmission of packets, even in traffic condition.. In a large network with high density, mobile nodes often create broken links. Maintenance and updating of tables as well as advertising would be significantly complex in this kind of network. DSDV is effective for ad hoc networks having small number of mobile hosts with limited changes in network topology.

1.2.2 Dynamic Source Routing (DSR)

DSR is a reactive protocol which reacts on-demand. The main feature of DSR is source routing in which the source always knows the complete route from source to destination. It frequently uses source routing and route caching. Route Discovery and Route Maintenance are the two main methods used in DSR. Any broken link is notified to the source node with an error message. It works well in large networks where routes change quickly and the mobility of routes is higher.

1.2.3 Temporally Ordered Routing Algorithm (TORA)

TORA is a highly adaptive loop-free distributed routing algorithm. TORA is proposed to operate in a highly dynamic mobile networking environment. It is source initiated and provides multiple routes for any desired source-destination pair. The key design concept of TORA is the localization of control messages to a very small set of nodes near the occurrence of a topological change. To accomplish this, nodes need to maintain routing information about adjacent (one-hop) nodes. The protocol performs three basic functions: route creation, route maintenance, and route erasure. It creates and maintains directed acyclic graph rooted at the destination node. TORA can establish routes rapidly and can provide multiple routes for a single destination. It does not give importance to the Shortest-Path algorithm. Instead, it uses longer paths to avoid finding of new routes. TORA minimizes communication as it reacts only when needed and doesn't react to every topological change.

1.2.4 Ad-Hoc On Demand Distance Vector (AODV)

AODV is a reactive protocol that reacts on demand. It is a modification of DSDV. The demand on available bandwidth is significantly less than other proactive protocols as AODV doesn't require global periodic advertisements. It enables multi-hop, self-starting and dynamic routing in MANETs. In networks with large number of mobile nodes, AODV is very efficient as it relies on dynamically-established route table entries at intermediate nodes.

1.2.5 Topology Broadcast Based on Reverse-Path-Forwarding (TBRPF)

TBRPF is a pro-active, link-state routing protocol for mobile ad hoc network which provides hop by hop routing along shortest paths to each destination. [4]It is useful when a large number of routes are needed and for applications that cannot tolerate the delay of route discovery. Each node running TBRPF computes a source tree to all reachable destinations based on partial topology information using modified

Dijkstra's algorithm. To minimize overhead, each node broadcasts only part of its source tree called the reportable tree to its neighbors. It uses a combination of periodic and differential updates to keep all the neighbors informed about the reported part of its source tree.

2 Proposed Work

An Intrusion Detection System (IDS) is a defense system that detects hostile activities in a network and then tries to prevent such activities that may compromise the system security.

The basic model for IDS proposed, runs an IDS agent on each node. It performs local data collection and local detection. The data collection module gathers local audit traces and activity logs that are used by the local detection engine to detect local anomaly. Both the local and global response modules provide intrusion response actions .The local response module triggers actions local to this mobile node whereas the global one co-ordinates actions among neighboring nodes.

Another kind of multi-layer IDS has also been proposed. [5]This system makes use of mobile agents which move around the network and work in parallel to execute their task. A mobile agent can cooperate with other agents, can copy itself and move from one platform to another, taking either a static or a dynamic route. The decision to move is made autonomously and by an independent active entity of the mobile agent and does not depend on the application which establishes it. MANET nodes have limited battery power, thus it is not efficient to make each node always a monitoring node. Instead, a cluster of neighboring MANET nodes can randomly and fairly elect a monitoring node, the cluster head, for the entire neighborhood i.e.the responsibility of intrusion detection is shared among all nodes in the cluster. As a result, cluster-based IDS schemes have been proposed.

Anomaly is also detected using cross-feature analysis, where various features are compared to other features. The features for classification may be either traffic related (e.g. packet-type, flow direction etc.) or non-traffic related (e.g. velocity, route cache etc.). Once the anomaly has been detected, the attack type and attacker is determined by getting the detailed attack information from a set of identification rules, which have been pre-computed for known attacks.

Illustration of Transient Loops:

An example of the evolution of a transient routing loop is shown in Fig.4 below. Consider a small network of three nodes-R1, R2, and R3 which are connected, with solid lines representing physical links, and dotted lines representing the flow of traffic between nodes. Nodes with consistent routing state are shown with the same shading. Traffic from all three nodes to other networks initially travels through R1 as depicted in Fig.4 (a); R2 has also advertised an alternative route. The link connecting R1 to other networks fail and R1 is the first to detect the failure. Traffic should now flow to other networks via R2, as shown in Fig.4 (d). However, until R2 learns of the failure of R1, it continues to direct traffic for other nodes in the network to R1. Since R1 knows that its link is broken and an alternative path is available via R2, it forwards this traffic back to R2 resulting in traffic looping as shown in Fig.4 (b). In Fig.4(c) the updated information has reached R3 before R2. Thus R3 begins sending traffic

destined for other networks to R2. As R2 is still unaware that R1 has failed, it continues to send this traffic to R1, which returns it to R2. Finally, in Fig.4 (d), information about link failure reaches R2 and so it stops routing traffic for other networks to R1, and uses its own link instead. Thus the loop finally disappears.

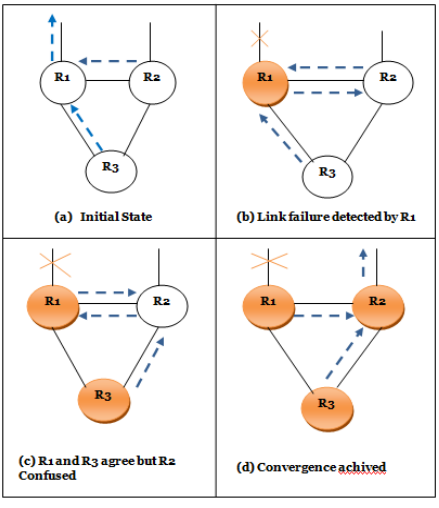


Fig. 4. Scenario for transient

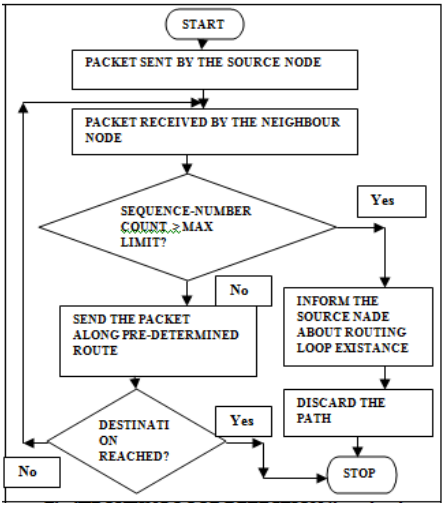


Fig. 5. Routing loop detection flow-chart Routing loop

3 Conclusion and Future Work

A MANET is a promising network technology, based on a self organized and rapidly deployed network. Access to information and services is regardless of geographic position. Networks can be set up at any time and they work without any pre-existing infrastructure. Due to great features, MANET attracts different real world application areas where the network topology changes very quickly. However, MANETs are more vulnerable to attacks than wired network due to their open medium. It is a dynamically changing network topology and hence mobile nodes come and go from the network, thereby allowing any malicious node to join the network without being detected. There is a lack of centralized monitoring and infrastructure which prohibits any monitoring agent in the system.

I have only discussed the security issues in this paper, particularly the routing loop attack.

References

1. Khokhar, R.H., Ngadi, A., Mandala, S.: A Review of Current
2. Routing Attacks in Mobile Ad Hoc Networks. International Journal of Computer Science and Security 2(3)

3. Lin, T., Midkiff, S.F.: Mobile Ad-hoc Network Routing Protocols. Chair Computer Engineering
4. Perkins, C.E., Bhagwat, P.: Highly Dynamic Destination-Sequenced Distance-Vector (DSDV) for Mobile Computers. In: Proc. ACM Conf. Communications Architectures and Protocols (SIGCOMM 1994), London, UK (1994)
5. Topology Broadcast Based on Reverse-Path Forwarding (TBRPF). Internet Engineering Task Force (IETF)
6. Ogier, R.G., Templin, F.L., Bellur, B., Lewis, M.G.: Draft (November 2002)
7. Royer, E.M., Toh, C.-K.: A Review of Current Routing Protocols for Ad-Hoc Mobile Wireless Networks. IEEE Personal Communications

Application of GIS Spatial Analysis Method in Landscape Planning and Design —A Case Study of Integrated Land-Use Suitability Analysis of Nanjing Zhongshan Scenic Area

Yue Xu

School of Geographic and Oceanographic Sciences,
Nanjing University
Educational Technology Center, Southeast University,
210096, Nanjing, China
chinaxuyue@qq.com

Abstract. Taking integrated land-use suitability analysis of Nanjing Zhongshan Scenic Area for example, the paper implements the application of GIS spatial analysis method in landscape planning and design via the comprehensive analysis and evaluation of multifactor. The process includes several steps such as collecting basic data, building DEM, and generating analytical diagram, etc.

Keywords: GIS; Multifactor; Suitability Analysis; Landscape planning and design.

1 Introduction

Landscape planning and design is the combination of technology and art, which needs full integration of technology and aesthetic principles, only in this way can planning designers finish outstanding landscape planning and design [1]. Currently, most domestic planners begin to use the information technology to aid planning and design, including dealing with documents and managing data using Microsoft Office software, aiding design and drawing using AutoCAD software, processing image using Photoshop software, making animation using 3Dmax software, etc [2-6]. And some planning designers make spatial analysis with software of GIS, there are also designers who conduct experiments of three-dimensional landscape simulation [7, 8]. But overall, the application of GIS in landscape planning and design is still at the initial stage [9].

The process of landscape planning and design includes basic research, program design, producing results, etc [10]. GIS can be applied to almost the entire process of landscape planning and design [11, 12]. But its outstanding applications mainly focus on in spatial data management, spatial analysis and visual expression three aspects, as shown in Fig. 1.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 67–73, 2011.

As the main features that GIS is different from other computer system, spatial analysis in landscape planning and design can be applied to terrain analysis, vegetation analysis, landscape visual analysis, traffic analysis, construction analysis, land suitability analysis and other aspects. Spatial analysis based on database can help designers for a rapid and scientific analysis, which makes designers more easily to form a clear awareness and understanding to basic current conditions for the design objective, and which provides the reasonable basis for planning and design work. The paper research the application of GIS spatial analysis method in landscape planning and design via the aggregate analysis and evaluation of multifactor taking integrated land-use suitability analysis of Nanjing Zhongshan Scenic Area for example.

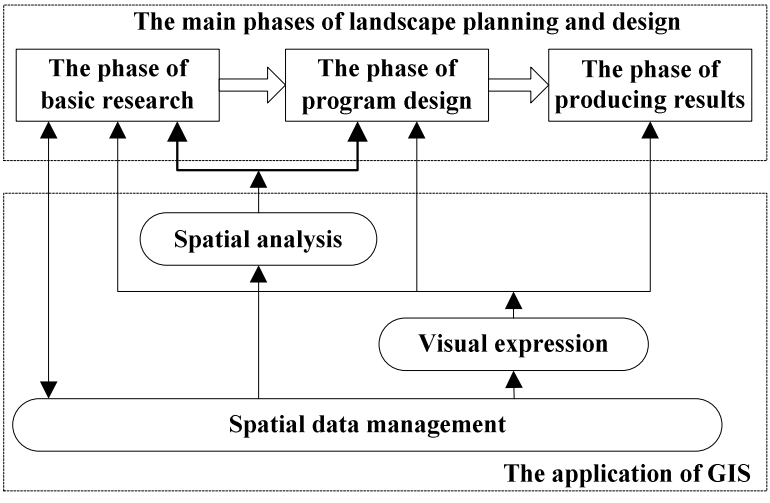


Fig. 1. The main phases of landscape planning and design and their application of GIS

2 Integrated Land-Use Suitability Analysis of Multifactor

2.1 Study Area

The Nanjing Zhongshan Scenic Area is located in the heart of urban development areas, with Rich natural ecological resources, profound historical and cultural resources and fair scenic resources. Recently, Nanjing Municipal Government attaches great importance to the planning and construction the scenic spot, conducts a comprehensive treatment surrounding environment, and establishes the development of cultural tourism and eco-oriented direction of future land use and development. With the successful completion of environmental comprehensive treatment Phase I, Phase II progressively plan and construct Philanthropy Park, Science Park, Sports Park, Folk

Garden, etc, in them, Philanthropy Park and Science Park are important parts of the core scenic spot, which are also the study area of the paper.

2.2 Technical Route

The study mainly includes collecting basic data, building DEM, and generating analytical diagram.

Collecting basic data. Landscape planning and design needing building spatial database and collecting the following two types of basic data:

a) Basic geographical data. The data mainly includes 1:10000 digital topographic map of the southern slope of Dr. Sun Yat-sen Mausoleum area and 1:500 digital topographic map of Philanthropy Park and Science Park. The maps mainly include terrain, buildings, roads and other surface features.

b) Various types of thematic data. The data mainly includes historic buildings (heritage) situation and vegetation coverage situation. The former includes building names, the existing situation, material, style, designers, geographical location, building age, reason of construction, building-type system, building type, cultural level, cultural relics, the scope of protection and other attribute data. The latter includes the forest area, the dominant tree species, the tree-age, the crown density, the forest physical characteristics, the forest structure, etc.

Building DEM. DEM is a digital expression of terrain surface and basic of spatial analysis and landscape mapping. The research adopts DEM which is generated by existing digital topographic map directly.

Generating analytical diagram. Based on DEM and spatial database, the research generates the required analytical diagrams via overlaying multiple factors participating in the evaluation according to certain rules.

3 Integrated Land-Use Suitability Analysis of Multifactor

3.1 Factor Selection

Land-use suitability analysis based on GIS is a comprehensive evaluation method of multifactor; therefore, the first step is to determining the factors participating in the evaluation. According to the simplicity (which means simple, easy to use and easy to calculate), scalability (its index values can be measured and accessed directly or indirectly, which is also quantitative.), and the main factors principles, select topography and vegetation, which are used as the factor of buildable analysis.

3.2 Comprehensive Evaluation and Analysis Process

When land-use suitability is evaluated, the research often needs to consider the combined effect of many factors. Based on determining the participating factor weight,

for each unit the level index of each index factor multiplied by the respective index weights, then accumulate products in order to the total score of comprehensive evaluation for each unit, finally, determining the buildable level for each unit according to high and low ranking of the total scores. Its computation formula is:

$$V(j) = \sum_{i=1}^{n} F_i W_i \tag{1}$$

Where $V(j)$ is the total score of the unit j, F_i and W_i are respectively the level index of each index factor and the respective index weights, n is number of the participating factors.

Slope analysis: The research makes terrain analysis for the study area. Based on DEM model established, slope analysis factor is divided into 6 classes: <5, 5-10, 10-15, 15-20, 20-25, >25, according to national requirements that construction land must be between 0-25 degrees. Slope map generated in ArcGIS is shown as Fig.2.

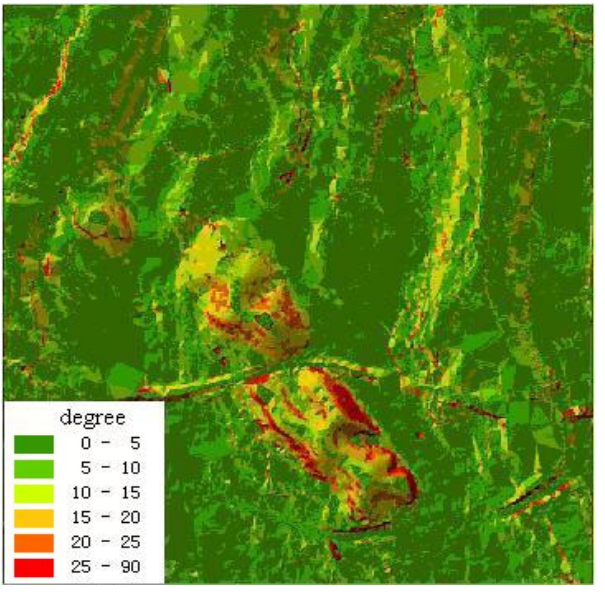


Fig. 2. Slope analysis diagram

Main evaluation factors: According to 1:10000 survey data for the forest in the study area, acquire the dominant tree species, the tree-age and the crown density of the forest. According to buildable suitability, grade the dominant tree species, the tree-age and the crown density in order to acquire their thematic analysis maps (Table 1, Fig. 3).

Table 1. Grading and weight of Evaluation factors

Evaluation factors	Evaluation content	The suitability's describe	Mark	Weighted value	
The dominant tree species	Evergreen coniferous forest	Weak ecological stability	1	0.5	0.5
	Shrubbery	Middling ecological stability	2		
	Broad-leaved forest, Coniferous forest, Deciduous forest	Good ecological stability	3		
	Oak, Liquidambar forest	Fairly good ecological stability	4		
Tree-age	Young forest, Over mature forest	Be suitable for construction	1	0.25	
	Half-mature forest	Be fairly suitable for construction	2		
	Near-mature forest	Be unsuitable for construction	3		
	Mature forest	Be awfully suitable for construction	4		
Crown density	<0.4	Bare woods, open forest	1	0.25	
	0.4-0.6	Open forest	2		
	0.6-0.8	Quite closed forest	3		
	>0.8	Closed forest	4		
Slope	<5%	Quite suitable	1	0.5	
	5%-15%	Suitable	2		
	15%-25%	Commonly suitable	3		
	>25%	Unsuitable	4		

Evaluation process: Firstly, make comprehensive evaluation of multifactor to the three vegetation single factor, acquire the analytical map for comprehensive evaluation of vegetation suitability (Fig. 4), in the map the higher the score is, the higher the comprehensive quality of vegetation is, and the area is more unsuitable for construction. Then, make comprehensive evaluation for vegetation and slope. Finally, acquire the analytical map for the buildable comprehensive evaluation of the study area (Fig. 5).

Evaluation result: The zone that acquires high score is unsuitable for constructing. According to the result of comprehensive evaluation, suitability of construction land is divided into four levels. Grade 4 areas are the highly ecological-sensitive areas which are prohibited for construction. Grade 3 areas are the middling ecological-sensitive areas which are unsuitable for construction. Grade 2 areas are the lowly ecological-sensitive areas which are conditionally-permitted for construction. Grade 1 areas are not the ecological-sensitive areas which are permitted for construction. Based on this, acquire land-use planning map of study area (Fig. 6), the planning buildings must mostly locate at the zone that is suitable for construction.

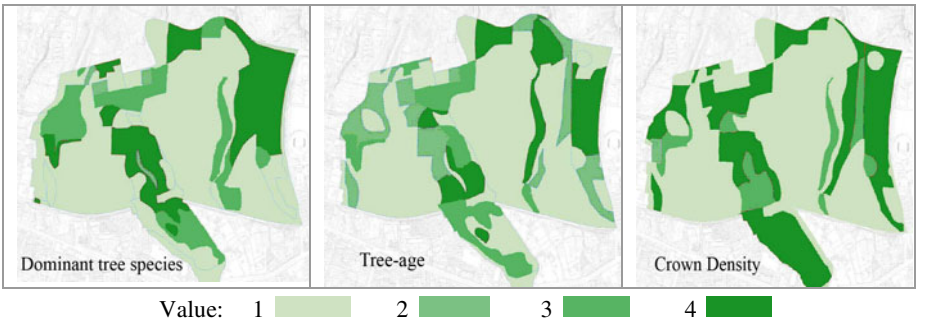


Fig. 3. Analysis diagram of vegetation's factors

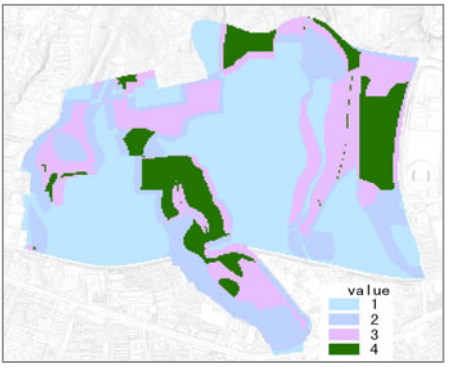


Fig. 4. Analytical map for comprehensive evaluation of vegetation suitability

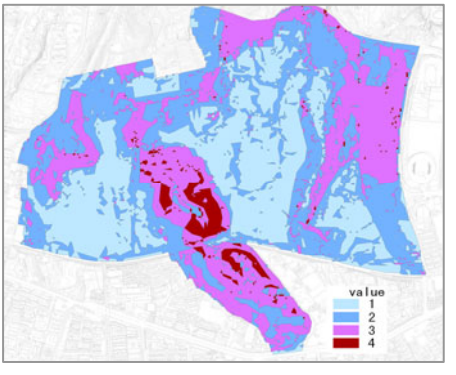


Fig. 5. Analytical map for the buildable comprehensive evaluation

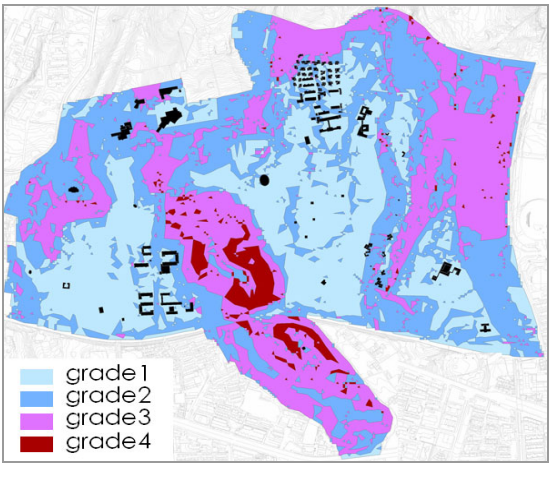


Fig. 6. Land-use planning map of study area

4 Conclusion

Data is the basis of the landscape planning and design. The quality of the planning depends largely on the quality of data collection and the scientificity of analytical approach. As a useful attempt to application of new technologies, the research enhances the scientificity of the planning, providing a powerful analytical tool and a good example for the follow-up landscape planning and design. However, new technologies is only a way to the planning, the more important is design with nature, and is to promote the harmonious development of Regional landscape, economy and humanity through the planning.

Acknowledgment. The author would like to express appreciations to Prof. Jianguo Wang & Associate Prof. Mingxing Hu of Southeast University for their valuable comments and other helps.

References

1. Donmez, S., Aciksoz, S.: Landscape Planning and Design Principles of Holiday Village: Antalya Kemer Ulusoy Holiday Village. African Journal of Agricultural Research 5(12), 1448–1455 (2010)
2. Wei, E., Lu, Y., Dong, X.: Residential Landscape Design of the Reconstruction of Villages in the City. Journal of Anhui Agricultural Sciences 38(3), 1640–1642 (2010)
3. Sun, X.: Ecology, Thrift and Legacy—Application of Earthy Landscape Elements in Urban Wetland Park Planning and Design. Journal of Nanjing Forestry University (Humanities and Social Sciences Edition) 9(04), 105–109 (2009)
4. Yin, C., Jiang, J.: Study on the Urban River Valley Landscape Design: The Lasa River Valley Landscape Design Example. J. Planners 2, 64–68 (2011)
5. Lu, Y., Li, S., Li, Y.: The Combination of the Landscape of City Park and Culture. Journal of Anhui Agricultural Sciences 39(5), 2966–2968 (2011)
6. Lang, W., Chen, J., Tian, G.: Landscape Planning and Ecological Restoration of Phoenix Mountain Forest Park in Xinxiang. Journal of Northwest Forestry University 26(1), 205–208 (2011)
7. Lange, E., Hehl, L.S.: Integrating 3D Visualization in Landscape Design and Environmental Planning. J. Gaia-Ecological Perspectives for Science and Society 15(3), 195–199 (2006)
8. Yu, L., Yu, Z., Pan, Y.: A Case Study of the Design and Application of 3D Visualization System For Agricultural Landscape Planning. Intelligent Automation and Soft Computing 16(6), 975–984 (2010)
9. Zhang, Y., Qiao, L.: The Application of GIS Technology in Urban Landscape Planning & Design in China. In: Proceedings, 2008 International Workshop on Education Technology and Training and 2008 International Workshop on Geoscience and Remote Sensing, pp. 95–98 (2009)
10. Liu, B.: On the Hot-points of Landscape and Design Domestically and Overseas: Theory, Technique and Creation. J. Urban Planning Overseas 2, 10–14 (1999)
11. Niu, X.: GIS Application in Urban Design. J. Urban Planning Forum 4, 41–45 (2002)
12. Song, X., Ye, J.: Geographic Information System and Its Application in Urban Planning and Management. Science Press, Beijing (1995)

Parameter Design for Operating Window Problems: An Example of Paper Feeder Design

Hsu-Hwa Chang[1,*], Jui-Yen Yen[1], and Tse-Chieh Lin[2]

[1] Department of Business Administration
National Taipei College of Business
No. 321, Sec. 1, Chi-Nan Rd., Taipei, Taiwan
[2] Lunghwa University of Science and Technology
Department of Industrial Management
hhchang@webmail.ntcb.edu.tw

Abstract. The operating window (OW) is the range between two performance limit thresholds if a system has a binary-type performance. Paper feeder design is a typical problem of the OW method. The wider OW, the higher performance of the system is. This study uses an artificial intelligent approach to optimize the OW design of a paper feeder. The approach employs an ANN to construct the response function model (RFM) of the OW system. A novel performance measure (PM) is developed to evaluate the OW responses. Through evaluating the PM of the predicted OW responses, the best control factor combination can be obtained by annealing simulated (SA) algorithm. An example of a paper feeder design is analyzed to confirm the effectiveness of the approach.

Keywords: Artificial Neural Networks, Operating Windows, Response Function Model, Annealing Simulated Algorithm, Paper Feeder Design.

1 Introduction

The concept of the operating window (OW) was developed by Clausing [3]. He used an OW response for the design of a friction-retard paper feeder in a copier machine. The function of a paper-feeding mechanism in a copier machine to feed exactly one sheet of paper each time the mechanism receives an input signal. When the mechanism does not feed any paper, it is called "misfeed." When two or more sheets of paper are fed into the copier machine at the same time, it is called "multifeed." This mechanism applies friction between the feeder roller and the paper, and the torque of the feed roller feeds the paper into the printer [4, 9, 11].

The friction force between the feed roller and the paper is determined by the spring force applied below the paper tray. When the spring force is too small, no paper will be sent out of the paper tray (misfeed). When the spring force is properly set, one sheet of paper will be sent out. When the spring force is set too large, two or more sheets might be sent out of the tray (mulitfeed). The objective the paper feeder design is to minimize the rate of both failure modes, i.e., misfeed and mulitfeed.

* Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 74–80, 2011.

Herein, spring force is a critical parameter of the paper feeder and is easy to measure. Let the threshold value of the spring force for sending one sheet of paper be x (gram-force). Let the threshold value of the spring force for multifeeding two or more sheets be y (gram-force). We can find two threshold values of the force at which the misfeed stops (x) and at which the multifeed starts (y). Then, (x, y) forms the OW [12, 15].

Thus, the objective of the paper feeder design becomes to minimize x to decrease number of misfeeds, and to maximize y to decrease number of multifeeds. Figure 1 shows two situations of an operating window. The operating window of situation B is wider and has a greater robustness than the window for situation A [16].

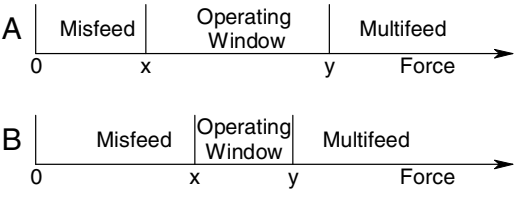


Fig. 1. Operating window

The two failure modes can be eliminated if x is reduces to zero and y is increase to infinity. Therefore, x is a small-the-better (STB) characteristic and y is a larger-the-better (LTB) characteristic. The optimization of the OW can be treated as to optimize simultaneously the responses of both STB and LTB in a system.

This paper aims to optimize the OW via parameter design. An artificial intelligent based approach is employed to here. First, an artificial neural network (ANN) is used to the system's response function model (RFM) that represents the relationship function between the OW responses and some control factors. The ANN can be applied to predict the OW responses by giving specific control factor combinations. Second, for evaluating the responses the OW responses are then integrated into a performance measure (PM) by exponential desirability functions. Finally, the annealing simulated (SA) algorithm is used to obtain the optimal control factor combination that has the maximum of the PM.

2 Exponential Desirability Functions

The exponential desirability function approach was introduced by Harrington [6] and further modified by Kim and Lin [10] and Chang [2]. The exponential desirability function transforms an estimated response (e.g. the $\hat{r}_j$ estimated response) to a scale-free value d_j, called desirability. It is a value between 0 and 1, and increases as the desirability of the corresponding response increases [8]. Goik et al. [5] firstly applied desirability functions to operating windows design. To evaluate different types of quality characteristics, the desirability functions are employed here and are slightly modified.

For the LTB type with lower specification limit (LSL), the desirability function of the d value (denoted by d^{LTB}) is formulated as Equations (1) and (2).

$$d^{LTB} = \exp(-(\exp(-Z^{LTB}))) \quad (1)$$

where

$$Z^{LTB} = \frac{\hat{r} - r_{\min}}{r_{\min}}, \quad (2)$$

$r_{\min}$ represents the LSL of response *r*.

For the STB type with upper specification limit (USL), the desirability function of the *d* value (denoted by d^{STB}) is formulated as Equations (3) and (4).

$$d^{STB} = \exp(-(1+Z^{STB})) \quad (3)$$

where

$$Z^{STB} = \frac{\hat{r} - r_{\max}}{r_{\max}}, \quad (4)$$

$r_{\max}$ represents the USL of response *r*.

3 ANN and SA

Recent works have discussed applying the method of integrating ANN and SA to optimize the parameter settings of engineering designs. The ANN is used to construct the relationship function between parameters and response, the SA is then applied to the network for searching the parameter settings with an optimal response. ANNs are composed of processing elements and connections. Among several networks the supervised learning network named back-propagation network (BPN) is most suitable for applying to parameter design because its ability of approximating any continuous mapping from the input patterns to the output patterns. BPN is a multi-layer network with learning ability. The nonlinear transfer function of sigmoid function $f(x) = 1/(1+e^{-x})$ is used between the connections of input layer, hidden layer, and output layer. Each layer is formed by several nodes and an additional bias node. BPN learning employs a gradient-descent algorithm to minimize the mean-square error (MSE) between the target data and the predictions of the neural network. The training data set is initially collected to develop a BPN model. Through a supervised learning rule the data set comprises of an input and an actual output (target). The gradient-descent learning algorithm enables a network to enhance its performance by self-learning. The training of a BPN involves three stages: the feedforward of the input training data, the calculation and back-propagation of the associated error, and the adjustment of the weights. While training the network model, the performance of the model is sensitive to various network structure choices and the parameter settings of learning rate and momentum coefficient. A common approach to obtain a well-trained network structure is to use the trial and error method, i.e., we can train several candidate networks that have a different number of hidden layers and nodes in each hidden layer, and then select the one with the smallest root of MSE (RMSE) [13, 14].

A SA algorithm is similar to local search with the added advantage of not being trapped in local optima. Starting from an initial solution, SA generates a new solution x' in the neighborhood of the current solution x. Then, calculate the change in the energy (objective) function, i.e., $\Delta E = f(x') - f(x)$. In maximization problems, if $\Delta E > 0$, transition to the new solution is accepted. If $\Delta E \le 0$, then transition to the new solution (i.e., an inferior solution) is accepted with a specified probability obtained by the function $e^{\Delta E/T}$, where T is a control parameter called the temperature. The initial value of the temperature parameter can be determined by using the function $e^{\Delta E/T}$. The T is gradually decreased by a cooling function $T^{new} = \alpha T^{old}$, where α is a cooling factor between 0 and 1. The probability of the acceptance of uphill moves decreases as T decreases. The SA algorithm needs to start from a high enough value of T so that all transitions are accepted. However, if this initial value of T is too high, it causes a waste of processing time. Applying the SA algorithm to a specific problem allows us to define the configuration of the possible solutions, the neighborhood of a solution, an energy function, and the annealing schedule. In addition, in this study the definition of the neighborhood refers to the *j*-neighborhood. The *j*-neighborhood of a solution vector x means selecting any j parameters and then randomly assigning another setting for each of them. Alternatively, the annealing schedule includes the parameter settings of T, L and α, and the stopping condition that terminates the algorithm. An effective cooling schedule is essential for reducing the amount of time required by the algorithm to find an optimal solution [2, 7].

4 The Approach

The proposed approach for treating the OW response problem comprises three phases. The first phase involves collecting experimental data for training an ANN to represent the RFM of the system, which is capable of predicting the corresponding OW responses by giving a specific factor combination. In the second phase, a novel performance measure derived from exponential desirability function is developed for evaluating the OW responses. The third and final phase provides the optimization processes to maximize the OW responses by using the RFM and the performance measure.

4.1 Phase 1: Response Function Modeling

This phase uses an ANN to model the response function. The input and output data are assigned as the level values for the control factor and the OW responses, respectively. A well-trained ANN represents the system's RFM. The process of the response function modeling consists of four steps, which are as follows:

Step 1. Randomly select the training and testing patterns from the experimental data.
Step 2. Select several ANN structures including all layer nodes.
Step 3. Set learning rate, momentum coefficient and executions iterations for each ANN structure.
Step 4. Train and choose a best ANN as the RFM.

4.2 Phase 2: Evaluating Performance Measures

For the paper feeder design, two OW responses (i.e., *x* and *y*) are simultaneously determined by of the system's control factor combinations. To measure the performance of the response *x* and *y*, the exponential desirability functions are employed here. For the response *x*, the desirability can be formulated as Equation (5).

$$d^x = \exp\left(-\left(1+\frac{\hat{x}-x_{\max}}{x_{\max}}\right)\right), \tag{5}$$

where $x_{\max}$ represents the USL of the OW response *x*, which is determined by the designer.

For the response *y*, the desirability can be formulated as Equation (6).

$$d^y = \exp\left(-\left(\exp\left(-\frac{\hat{y}-y_{\min}}{y_{\min}}\right)\right)\right), \tag{6}$$

where $y_{\min}$ represents the LSL of the OW response *y*, which is determined by the designer.

To measure the overall performance of the paper feeder system, two OW responses need to be integrated into a single performance measure (denoted by *PM*). To enhance the overall performance PM, the optimizing of the OW response problem can be stated as:

$$\text{Maxminize} \quad PM = \sqrt{d^x \cdot d^y} \tag{7}$$

4.3 Phase 3: Obtaining Optimum

The phase uses a SA algorithm to perform the optimization process:

Step 1. Set the parameters of the SA including the temperature, cooling factor and the terminated condition. Define the objective function (i.e., *PM* value).
Step 2. Perform the SA Algorithm.
Step 3. Obtain the best PM value and the corresponding control factor combination.

5 Implementation

An example of paper feeder design adopted from Chang and Chen [1] is reanalyzed by the approach. Six control factors, A, B, C, D, E and F, are selected and are allocated in the L_{18} orthogonal array (OA) for the experiments. Table 1 lists the control factor levels and their allocations. The experimental data including misfeed threshold (*x*) and multifeed threshold (*y*) are listed in Table 2. The USL and LSL for the thresholds *x* and *y* are set as 500 and 400 grams, respectively.

The RFM can be built through training an ANN model. The ANN is trained by assigning the levels of control factors and the values of thresholds (i.e., *x* and *y*) as the inputs and outputs of the network. Eight patterns are randomly selected for testing and 64 patterns are selected for training. After training several networks, the networks structure 6-8-2 with the lowest testing RMSE, 0.1468, is chosen to obtain a better performance.

Table 1. The control factors and their allocations

Label	Factors	Column in L_{18}	Levels		
			1	2	3
A	Pad coefficient of friction	1	Low	High	-
B	Retard pad force	4	Low	Nominal	High
C	Retard angle (degree)	5	19	21	23
D	Feed roll to pad lateral offset	6	-2mm	Centered	+2mm
E	Width of feed belt (mm)	7	10	20	30
F	Roll velocity	8	Low	Nominal	high

Table 2. The experimental data

Exp. No.	*x* value (gram)				*y* value (gram)			
1	335	340	298	326	633	680	816	720
2	309	321	282	279	635	595	735	637
3	335	286	373	228	664	677	774	756
4	286	429	414	300	660	682	594	729
5	463	309	352	314	586	788	613	604
6	267	323	339	259	754	745	702	678
7	331	290	335	249	586	709	685	533
8	302	272	395	269	798	691	712	778
9	250	337	335	368	613	669	591	665
10	390	370	384	202	531	508	805	758
11	255	282	277	326	702	666	704	654
12	245	381	329	325	631	698	592	609
13	323	247	326	321	680	655	605	727
14	273	247	340	354	698	755	691	724
15	360	153	282	292	648	700	782	696
16	231	226	335	221	529	698	640	539
17	173	273	377	223	560	587	797	714
18	199	307	323	285	613	621	806	753

In the optimization process, the Equation (7) is employed to be the objective function of the SA. The SA's operating conditions including initial temperature, stopping temperature, cooling factor, and epoch length are set as 4, 0.001, 60, 0.95, respectively. The result is listed in Table 3 and is compared to the literature. The result shows that the approach has higher performance than the literature.

Table 3. The result and comparison

Method	Control factor settings						$\hat{x}$	$\hat{y}$	*PM* value
	A	B	C	D	E	F			
Chang and Chen [1]	High	High	19	+2mm	30	Nominal	216	691	0.6328
The approach	High	Low	19	+1mm	12	Nominal	220	709	0.6370

6 Conclusion

In this study, an artificial intelligent approach is proposed to resolving the parameter design of the operating window problems. A paper feeder design is analyzed by the approach to verify the approach's effectiveness. Performing the approach, engineers do not require much background in statistics but instead rely on their knowledge of engineering. The proposed approach can be also applied to other industrial systems that have binary-type performance such as wave soldering and resistance welding.

References

1. Chang, H.H., Chen, C.H.: Use of response surface methodology and exponential desirability functions to paper feeder design. WSEAS Transactions on Systems 7(1), 17–23 (2008)
2. Chang, H.H.: A data mining approach to dynamic multiple responses in Taguchi experimental design. Expert Systems with Applications 35, 1095–1103 (2008)
3. Clausing, D.P.: Operating window: an engineering measure for robustness. Technometrics 46(1), 25–29 (2004)
4. Fowlkes, W.Y., Creveling, C.M.: Engineering Methods for Robust Product Design. Addison-Wesley, Reading (1995)
5. Goik, P., Liddy, J.W., Taam, W.: Use of desirability functions to determine operating windows for new product designs. Quality Engineering 7(2), 267–276 (1994-1995)
6. Harrington, E.C.: The desirability function. Industrial Quality Control 21, 494–498 (1965)
7. Hsieh, K.L.: Process parameter optimization via data mining technique. In: Proceedings of the 6th WSEAS International Conference on Simulation, Modelling and Optimization, pp. 130–134 (2006)
8. Hsu, C.M., Su, C.T., Liao, D.: Simultaneous of the broadband tap coupler optical performance based on neural networks and exponential desirability functions. International Journal of Advanced Manufacturing Technology 23, 896–902 (2004)
9. Joseph, V.R., Wu, C.F.J.: Operating window experiments: a novel approach to quality improvement. Journal of Quality Technology 34(4), 345–354 (2002)
10. Kim, K.J., Lin, D.K.J.: Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability functions. Applied Statistics (Journal of the Royal Statistical Society: Series C) 49(3), 311–325 (2000)
11. Maghsoodloo, S., Ozdemir, G., Jordan, V., Huang, C.H.: Strengths and limitations of Taguchi's contributions to quality, manufacturing, and process engineering. Journal of Manufacturing Systems 23(2), 73–126 (2004)
12. Park, S.H.: Robust Design and Analysis for Quality Engineering. Chapman & Hall, Lodon (1996)
13. Pricipe, J.C., Euliano, N.R., Lefebvre, W.C.: Neural and Adaptive systems. John Weily & Sons, Inc., West Sussex (2000)
14. Rowlands, H., Packianather, M.S., Oztemel, E.: Using artificial neural networks for experimental design in off-line quality. Journal of Systems Engineering 6, 46–59 (1996)
15. Samuel, N., LaVallee, L.: Application of Taguchi method to friction retard feeder design. In: Case Studies & Tutorials: 8th Taguchi Symposium, ASI, pp. 335–372 (1990)
16. Taguchi, G.: Taguchi on Robust Technology Development. AMSE Press, New York (1993)
17. Wu, Y., Wu, A.: Taguchi Methods for Robust Design. ASME Press (2000)

Simulation of Atomic Clock Noise by Computer

Yuli Li* and Hongwei Sun

Xi'an University of Arts and Science,
168 South Taibai Road, Xi'an, 710065, Shaanxi, China
Sunhongwei100@yahoo.com.cn

Abstract. A method for simulating three types of noise in atomic clock is presented. For each noise type, one obtains the noise level from a plot of the conventional Allan variance, and we select a value to be the most reliable indication of the Allan variance, the value is taken as inputs to digital filter that is a specific recursive function to produce the random processes. The simulation is simple by computer, and it is useful for designing atomic clock before committing the system to a hardware realization.

Keywords: Simulation, Atomic clock, Noise.

1 Introduction

Typical noise model of atomic clock is a superposition of five noise types, they are named as white noise phase modulation (W-PM), flicker noise phase modulation (F-PM), white noise frequency modulation (W-FM), flicker noise frequency modulation (F-FM), and random walk noise frequency modulation (RW-FM) [1], and the model is called power-law spectra noise model [2]. For W-PM, W-FM, and RW-FM, we give a simulation method.

2 Simulation Method

Allan variance is expressed as

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2} E[(x_n - 2x_{n+N} + x_{n+2N})^2], \tag{1}$$

where x_n is time datum, $\tau = N\tau_m$, τ_m is the interval between points [3].

2.1 Simulation to W-PM

Let a_n denote the random number of normal distribution, and their mean is zero and the variance is σ_a^2, let x_n denote the random number to simulate W-PM. First of all, it is need to calculate the variance of x_n.

* According to the Chinese naming convention, the name should be Li Yuli. This work was supported by Grant 2010JK796 and 2010JK799 from the Department of Education of Shaanxi in China.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 81–85, 2011.

For the phase white noise, have $x_n = a_n$, substitution it into (1), have

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2} E[(a_n - 2a_{n+N} + a_{n+2N})^2] \cdot \tag{2}$$

Based on

$$E(a_n, a_m) = \begin{cases} 0, n \neq m \\ \sigma_a^2, n = m \end{cases},$$

we obtain

$$E[(a_n - 2a_{n+N} + a_{n+2N})^2] = 6\sigma_a^2, \tag{3}$$

that is

$$\sigma_a^2 = \frac{1}{3}\tau^2 \sigma_y^2(\tau) \cdot \tag{4}$$

The noise level is from the Allan variance of conventional atomic clock. Based on known Allan variance to be simulated over a suitable range of time intervals, we select a point (τ_{-2}) which appears typical of the noise type to be simulated, and $\sigma_y^2(\tau_{-2})$ is the corresponding Allan variance, thus, we obtain the variance of the random numbers to simulate the noise

$$\sigma_{-2}^2 = \frac{1}{3}\tau_{-2}^2 \sigma_y^2(\tau_{-2}) \cdot \tag{5}$$

Then, random numbers to simulate the desired W-PM are

$$x_n = \sigma_{-2} p_n, \qquad (n = 1, 2, \ldots) \tag{6}$$

where p_n is random number with standard normal distribution. Fig.1 shows the simulation data to W-PM.

2.2 Simulation to W-FM

Let y_n denote the random number to simulate W-FM, thus $y_n = a_n$. In this case, the variations in frequency affect the rate of change of phase variations. For simple, let the interval of sample be one unit and x_0 be zero, thus

$$y_n = x_n - x_{n-1} = a_n,$$

that is

$$x_n = \sum_{i=1}^{n} a_n \cdot \tag{7}$$

Substitution it into (1), have

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2} E[(\sum_{i=1}^{n} a_i - 2\sum_{i=1}^{n+N} a_i + \sum_{i=1}^{n+2N} a_i)^2]$$

$$= \frac{1}{2\tau^2} E[(\sum_{i=n+N+1}^{n+2N} a_i)^2 + (\sum_{i=1+n}^{n+N} a_i)^2]$$

$$= \frac{N}{\tau^2}\sigma_a^2,$$

thus

$$\sigma_a^2 = \frac{\tau^2}{N}\sigma_y^2 \cdot \tag{8}$$

On the noise level, we select a point (τ_{-1}) which appears typical of the noise type to be simulated, and $\sigma_y^2(\tau_{-1})$ is the corresponding Allan variance. Thus, we obtain the variance of the random numbers to simulate the noise

$$\sigma_{-1}^2 = \frac{\tau^2}{N}\sigma_y^2(\tau_{-1}), \tag{9}$$

thus, random numbers to simulate the desired W-FM are

$$x_n = x_{n-1} + \sigma_{-1} p_n \cdot \qquad (\text{n = 1, 2, } \dots) \tag{10}$$

Fig.2 shows the simulation phase data to W-FM.

2.3 Simulation to RW-FM

Let z_n denote the random number to simulate RW-FM, thus $z_n = a_n$. Since the variations affect the rate of change of frequency variation, we obtain

$$z_n = y_n - y_{n-1} = a_n,$$

that is

$$y_n = \sum_{i=1}^{n} a_n \cdot$$

Because of $y_n = x_n - x_{n-1}$, have

$$x_n = \sum_{j=1}^{n}\sum_{i=1}^{j} a_n$$

$$= \sum_{i=1}^{n}(n-i+1)a_i \cdot$$

Substitution it into (1), have

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2}E[(\sum_{i=1}^{n}(n-i+1)a_i - 2\sum_{i=1}^{n+N}(n+N-i+1)a_i + \sum_{i=1}^{n+2N}(n+2N-i+1)a_i)^2]$$

$$= \frac{1}{2\tau^2}E[(-\sum_{i=1+n}^{n+N}(n-i+1)a_i + \sum_{i=n+N-1}^{n+2N}(n+2N-i+1)a_i)^2]$$

$$= \frac{\sigma_a^2}{2\tau^2}[N^2 + 2(N-1)^2 + 2(N-2)^2 + \cdots + 2\times 1]$$

$$= \frac{\sigma_a^2}{6\tau^2}(2N^2+1)N^2 \cdot$$

We select a point (τ_1) which appears typical of the noise type to be simulated, and $\sigma_y^2(\tau_1)$ is the corresponding Allan variance. Thus, we obtain the variance of the random numbers to simulate the noise

$$\sigma_1^2 = \frac{\sigma_y^2(\tau_1)}{(2N^2+1)N^2}6\tau^2 \cdot \tag{11}$$

Random numbers to simulate the desired RW-FM are

$$z_n = \sigma_1 p_n .$$

That is

$$y_n = y_{n-1} + \sigma_1 p_n ,$$

or

$$x_n = 2x_{n-1} - x_{n-2} + \sigma_1 p_n . \qquad (n = 1, 2, \dots) \qquad (12)$$

Fig.3 shows the simulation phase data to RW-FM.

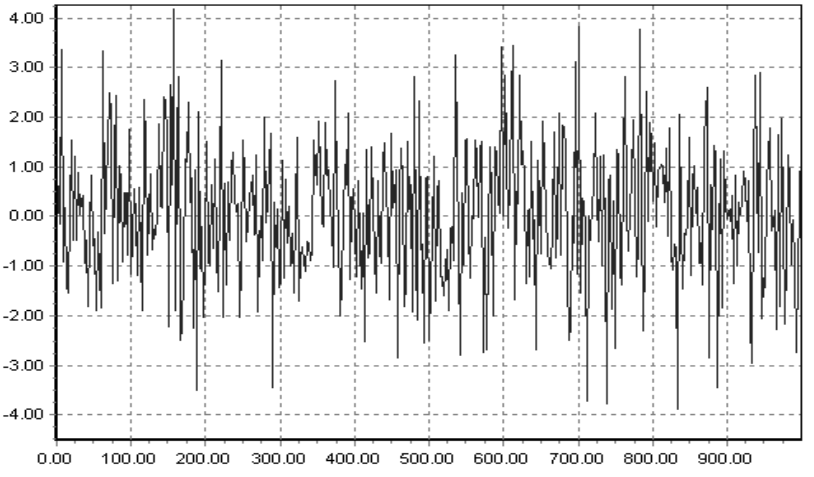


Fig. 1. Simulation data to white noise phase modulation (ordinate unit: s)

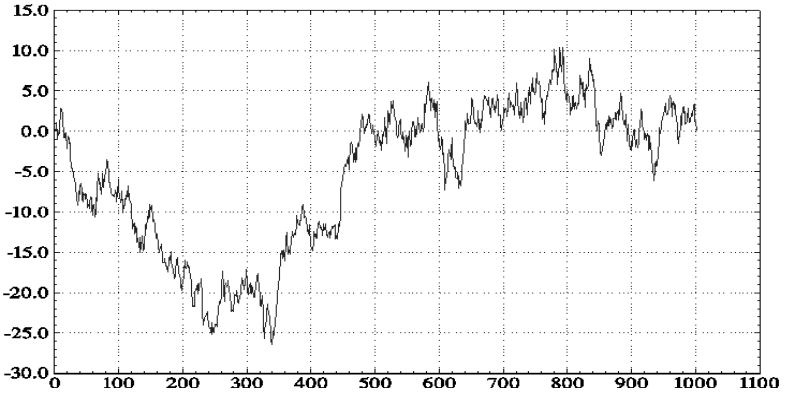


Fig. 2. Simulation phase data to white noise frequency modulation (ordinate unit: s)

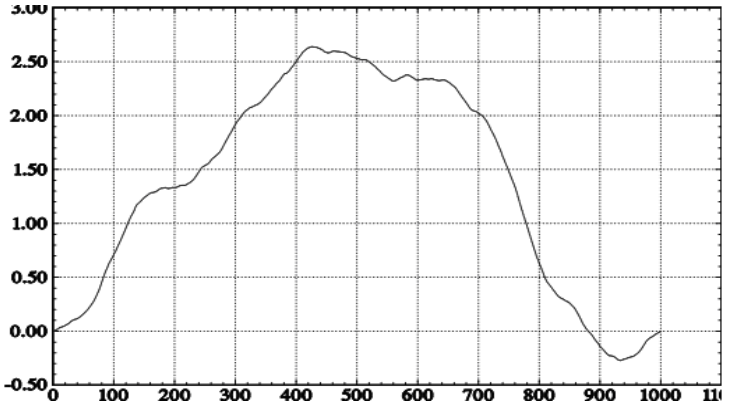


Fig. 3. Simulation phase data to random walk noise frequency modulation (ordinate unit: 100s)

3 Conclusions

Via the simulation method above, one can simulate the three types of noise in atomic clock. An Allan variance can indicate the levels and types of present noise. For a given noise type, we select a value to be the most reliable indication of the Allan variance, as a result, the recursive functions produce these random processes.

References

1. Barnes, A.: Characterization of Frequency Stability. IEEE Transactions on Instrumentation and Measurement 21, 11–20 (1997)
2. Schmidt, S.: Atomic Clock Models Using Fractionally Integrated Noise Processes. Metrologia 40, 305–311 (2003)
3. Allan, W.: Time and Frequency (time-domain) Characterization, Estimation, and Prediction of Precision Clocks and Oscillators. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 34, 647–654 (1987)

The Research on Remote Sensing Image Data Sharing Model Based on SOA

Lianbei Wang, Xin Wang, and Tao Wang

Institute of Surveying and Mapping, Information Engineering University,
Zhengzhou 450052, China
{wlbzhm,wangxin,wangtao}@gmail.com

Abstract. Based upon the in-depth analysis of the remote sensing image sharing demand and the status of the current research on it, this paper points out the existing Web-based sharing-model did not eliminate "Information Islands" fundamentally. To tackle the problem, the author has been working on the research topic of finding a remote sensing image data sharing model based on the service-oriented modern network computing theories and their supporting technology system. Hereby the author has provided the data sharing model based on SOA. He had realized the service-oriented RS image data sharing prototype system, and carried out the experiments on the prototype system. The result of the experiment shows, the service-oriented data sharing model is technologically available for the integration of heterogeneous and geographically distributed image storage and management system.

Keywords: RS image, Data sharing, SOA, Data service, Information Islands.

1 Introduction

The spatial geography information with remote sensing image as the main part has been widely employed in the surveying and mapping, the agriculture and forestry, and so on. For a long time however, the “Isolated Information Islands” have resulted from no interconnection among the image data management systems built by different departments adopting different technical patterns. And the due utility of the image can not be brought into full play, which, objectively, has caused grave waste. Accordingly, how to share the heterogeneous RS image data resources distributed in the geography has become an issue desiderating to be solved urgently[1][2].

2 Research Status and Analysis

Because the data is characteristic of their different formats, heterogeneous storage and management system, geographical distribution, etc., the spatial data sharing is more challenging compared with the common data sharing. As for this question, it is solved traditionally by means of “unifying and transforming”, that is, unifying the data format and constructing the data storage and the management system on the basis of the unified technical pattern. But the way needing huge expenditure to transform and

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 86–92, 2011.

reconstruct the existing heterogeneity and heterogeneous data sources, the existing investment is unable to be protected and the data sharing cost is to be increased. In addition, it is a static heterogeneous elimination way and difficult to adjust to the technical and application dynamic development change. The change of the "transforming" standard will result in reconstructing the "unified" system and more importantly, the way has neglected the objectivity and the rationality of the existence of the heterogeneous data sources.

Concerning the technical means of the data sharing, some RS image issuing systems have appeared one after another based on the Internet along with the development of the network technology and the representatives are Microsoft TerraServer, Google Earth, NASA World wind and so on. It is known that there are mainly two kinds of data sharing patterns through analyzing the characteristics and the technical essence of the existing RS image issuing systems based on the Internet technology:

"Search – order" model: a user gets the address of the image data issuing website through the search engine, inquires the demand data through browsing the metadata and then gains the data through physical media like CD, etc. or downloading from the website. Its technology, actually, is to realize the metadata information issuing with the aid of the Web technology and provide the data downloading service through the protocols like HTTP, FTP, etc.. This is the most common image data sharing pattern at present, such as Microsoft TerraServer, China special information website and so on. The Internet search engine based on the homepage search cannot offer inquiring the dataset metadata, and accordingly the efficiency of the manual off-line sharing pattern and the inquiry accuracy ratio are all low.

"On-line browse" model: the user uses the special client (or browser plug-in) to inquire and browse on line the image data issued by the data provider, Google Earth, NASA World Wind, etc., for instance. Its technical essential is that the client uses the specific protocol to visit the image data management system of its server to gain the image data.

It is easy to find that the above sharing model doesn't truly eliminate the "isolated information island". It is mainly because from the technological analysis the Internet technology can solve the problem of the information issuing, but cannot well solve the problem of the data sharing and the reason is that in the existing Internet technology, there is still lack of an environment in which the data are easily shared to solve the related questions of sharing data among the distributed heterogeneous systems and the incompatible data access method is still used among various kinds of data issuing systems[3][4]. Along with the building of various data browsing and issuing systems based on different technical systems, the "isolated information island" not only has not been eliminated, which results in the new "isolated information island" instead.

The application demand variety decides that data sharing should essentially pursue "taking what one needs". The author holds that the ideal data sharing model should be the "Machine-to-machine on-demand Service" model under the modern information technology conditions, namely, the client application using data dynamically search the data service which is able to provide the data needed and directly gains the data demanded through the data service. The data sharing process should be carried out automatically and dynamically as far as possible; the manual participation should be reduced to enhance the sharing efficiency.

So far however, this model has not been applied universally. From the technological analysis, it is mainly because the traditional Internet technology cannot directly provide a convenient and simple approach for the owners of the RS data, which enables the owners to offer at a relatively low price the on-line, stable and reliable data access service for users in the WAN (for example Internet). On the other hand, the traditional Internet technology, also, cannot directly provide the effective means of finding data and integrating data service for the client application.

The appearance of the new-generation Internet computation thought and the realization technology, such as SOA, Web Services, Grid computing, etc., opens the door to solve the above issues[5][6].

3 The RS Image Data Sharing Model Based on The SOA

Aiming at the social and popular share requirement of the remote sensing image data, the paper discussed the relevant theories and the key technology issues involved in realizing the data share model of "machine-to-machine on-demand service" based on the service-oriented modern network computing theories and their supporting technology system. Based on SOA, the author designed a service-oriented distributed RS image data sharing model in the WAN environment supported by Web Services and the data grid technology. The model is composed of four layers (figure 1).

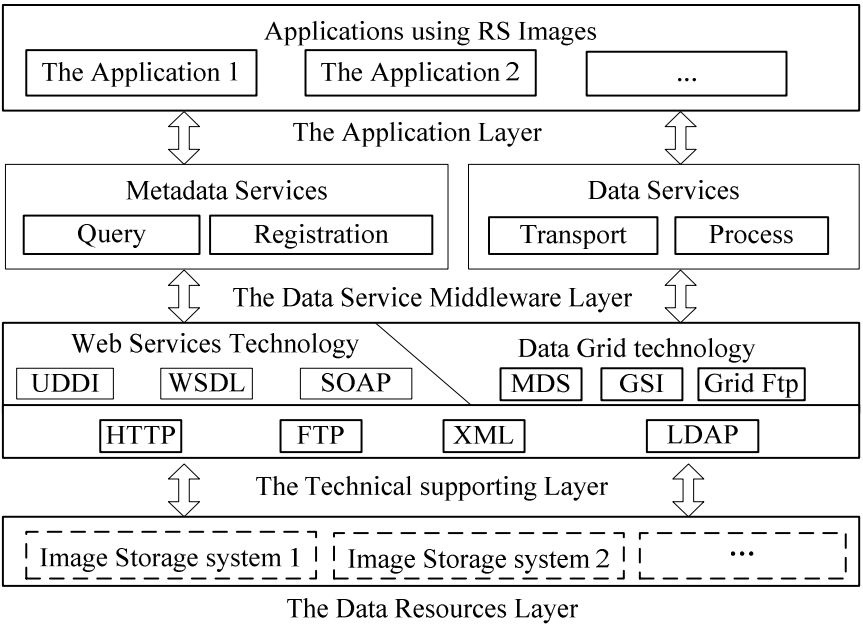


Fig. 1. The RS Image Data Sharing Model

The Application Layer: It is composed of various client programs which use the remote sensing image. It is the data user in the architecture.

The Data Service Middleware Layer: It is composed of the metadata registration and updating service component, the data query component and the data service component. It is the interface for the data users to find and acquire the data as well as the agent for the data providers to provide their data.

The metadata registration and updating service components are provided for the data providers to register and update the relevant metadata of the data contents and the data service.

The metadata query service component is provided for the client application to find the data which needed.

The data service components, which is composed of data transport service and data process service, is provided for the client application to process and get the data.

The Technical Supporting Layer: It is composed of the web service technology, internet technology, data grid technology. It provides the basic network computing environment to metadata and data service components in the upper layer.

The Data Resource Layer: It is composed of various remote sensing image storage and management systems. The systems might be the heterogeneous systems which are geographically distributed. They are the data providers in the model.

The model provides the highly effective data issuance and the discovery mechanism and the flexible integration mechanism of the heterogeneous system which enables the client application which is based on the standards interrelated with the space information field and the information technology domain finds and integrates the distributed heterogeneous image data management system, and then makes the sharing of the image data resources realized.

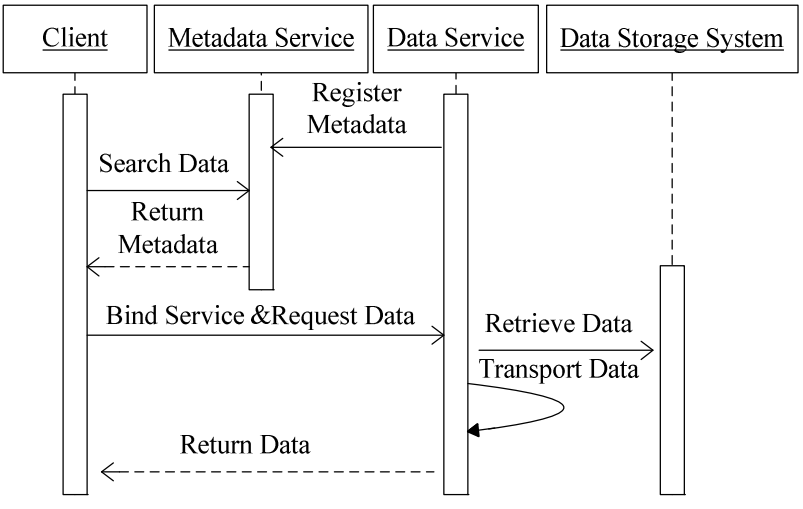


Fig. 2. The RS Image Data Sharing Interaction Pattern

In this model, the remote sensing image data sharing service works like this:

(1) The data providers register and update the metadata on the metadata service system.

(2) The client application searches the data via the metadata service system.

(3) The metadata service system finds out the data provider catering to the requirements and provides the information to the applicant.

(4) The client program binds the data service interface provided by the data provider and request data.

(5) The data service interface acquires the data from the data storage system and send it to the client application.

4 Experiments

In order to confirm the feasibility of the theories and technology method interrelated with the service-oriented image data sharing brought forward in the paper, the metadata service subsystem and the data service subsystem have been developed and the service-oriented RS image data sharing prototype system has been realized. The data sharing experiment platform consisting of the client application node, the metadata service node, the data service node as well as the image storage and management node has been built. On the basis of the experiment platform (figure 3), the data search and automatic acquisition experiment were made respectively.

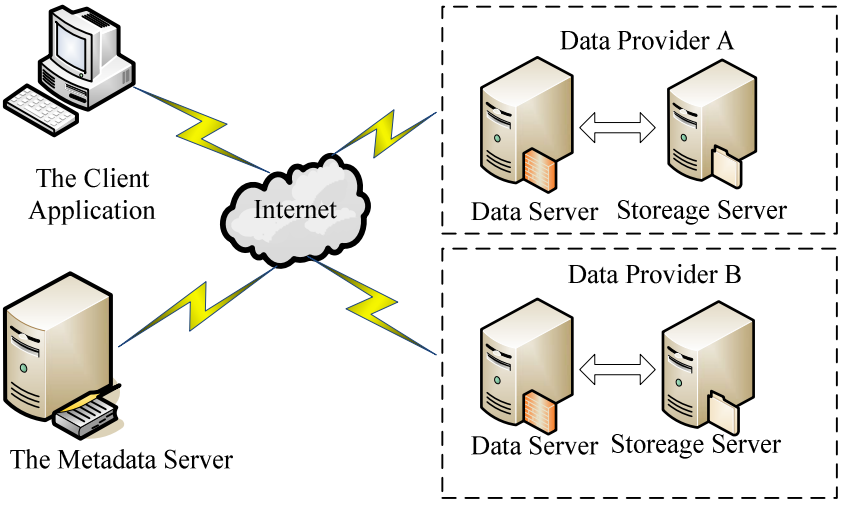


Fig. 3. Experimental Environment

In the experiment, the client application program can find out the appropriate data provider via the Metadata Management Services and bind the data services provided by the data provider, then get data. The Interface of the client application program is shown in figure 4.

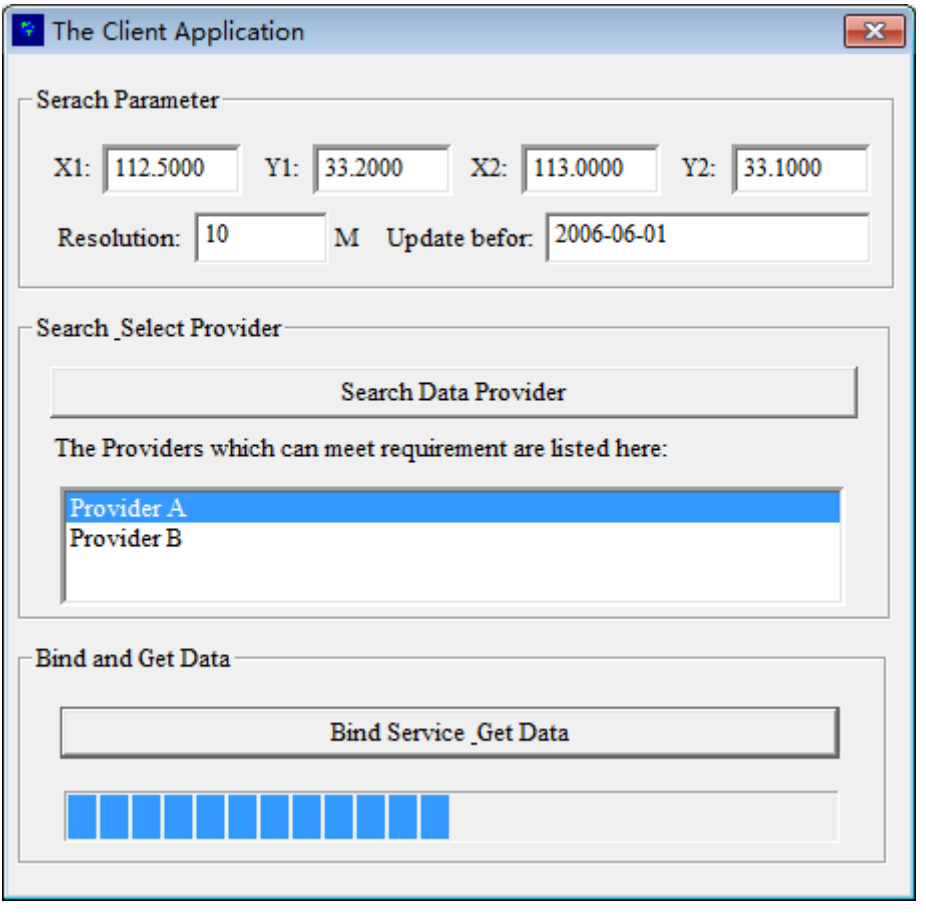


Fig. 4. The Interface of the client application program

The experiments indicate the image data sharing mechanism of the "machine-to-machine on-demand service" model is technically feasible and can flexibly realize the on-demand sharing of the distributed heterogeneous RS image data resources.

5 Conclusions

The socialized and popularized share of the RS image is a complex systems engineering. On the basis of understanding the connotation of the data share and the latest development of the information technology, the dissertation probed into the "machine-to-machine on-demand service" data sharing model and its technical realization method from the angle of the service-oriented. And the results of the study enrich the RS image data sharing theories and the technical realization method. The further study will be done encircling strengthening the expressive power of the

metadata model, extending the inquiry way of the image data, enhancing the intellectualized service discovery level of the data and so the safety mechanism needed in the data sharing process, etc.

References

1. Kim, T.J.: Metadata for geo-spatial data sharing: A comparative analysis. J. The Annals of Reglonal Seienee 33, 171–181 (1999)
2. Stoimenov: An architecture for interoperable GIS use in a local community environment. J. Computers & Geosciences 31, 211–220 (2005)
3. ErI, T.: Service-Oriented Architecture Concepts, Technology, and Design. Prentice Hall Ptr, New Jersey (2005)
4. Foster, I.: The Grid: A New Infrastructure for 21st Century Science. J. Physics Today 55, 42–47 (2002)
5. Huber, M., Schneider, D.: Spatial data standards in view of models of space and the functions operating on them. J. Computers & Geosciences 25, 25–38 (1999)
6. Map Server. University of Minnesota, `http://mapserver.gis.umn.edu`

Applying Problem-Based Learning in Course Design Teaching

Yumin Xie

Nanjing Institute of Technology
Nanjing, China
Xxyymm3721@126.com

Abstract. Puzzle-based learning, problem-based and project-based learning are main methods used by academics. When a teacher devises a course design plan with about half month duration, a puzzle-based plan may be too narrow while a projected-based plan may be too big at the same time. We talk about applying problem-based learning in course design in the paper. A winged-edge data structure is employed in our java course design plan in order to express a complex problem with a single data structure. With this complex problem, the plan can make students have great leeway to include as many as possible concepts, such as class, interface, encapsulation to their final work. Students are grouped in order to develop their team spirit in studying.

Keywords: PBL, Winged-edge data structure, Teaching research, Java, OO.

1 Introduction

In our school, programming JAVA is a course followed by a Course Design when the lesson finished. The design duration is about half month, which includes a week-long laboratory phase and a week-long self-design phase. As we all know that effective teaching depends on effective planning and design, so an appropriate plan should be work out first.

Three problems must be considered when a teaching staff to design such a plan:

1. Students' ability
2. Learning goals
3. Learning methods.

First, we teach programming JAVA course to Information Systems (IS) students other than Computer Science (CS) students. Unlike CS students, IS students are generally short of background or prerequisite courses when they progress a course like JAVA.

Second, two goals included in our plan according to course outline. One is the design should reflect the students level of their OO concepts. The other is to develop their team spirit, as we know that OO paradigm comes from the writing of large programs which involve several programmers.

In fact, when a teaching staff begins to consider such a plan, the students' ability and the goal are fixed. The staff only has leeway to select a learning method.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 93–99, 2011.

Now we review three main learning methods as below:

Project based learning[1] is a pedagogical approach that significantly increases the assimilation of knowledge by forcing students to solve problem solving opportunities within the framework of a complex, open-ended project. A primary characteristic of this approach is that it is less structured than traditional, faculty-led activities and problem-based learning - students often required to organize their own work and manage their own time.

Problem-based learning[1]is typically completed with small groups, all lead by a facilitator. A constructed but realistic problem is presented to the group. This problem may be in terminal form or maybe an enabling issue. An enabling problem is presented, students must engage in inquiry to get information about the terminal problem. Usually, "problems" are ill-structured and open to differing approaches and offer thematic sidelines. The students discuss the presented facts and data, further define problems, analyze the support data and make recommendations for solution. Typically, analysis and recommendations are presented in a formal format but non-formal presentations work well too. The process concludes with a debriefing which includes reflections on learning and lessons learned.

The puzzle-based learning[2] approach aims to encourage students to think about how they frame and solve problems not encountered at the end of some textbook chapter. This approach is very useful when we want to deep thinking a concept.

2 The Winged-Edged Data Structure [3]

In Fig. 1, edge a has incident faces 1 and 2, and the traversal of each incident face under the predefined orientation induces a predecessor edge and a successor edge. The predecessor and successor edges of edge a with respect to face 1 are edges b and d, respectively, and the predecessor and successor edges of edge a with respect to face 2 are edges e and c, respectively. Since each face induces a different direction to the common edge, a direction must be chosen in order to specify the "left" face and the "right" face. For example, in Fig. 1, if the chosen direction is from vertex X to vertex Y , the left face is face 1 and the right face is face 2. Otherwise, the left and right faces are 2 and 1, respectively.

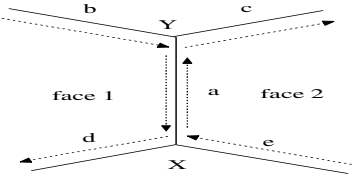


Fig. 1. Winged-Edged Data

Once the orientation of an edge is chosen, one can assemble nine pieces of information from an edge: the edge name, the start and end vertices, the left and right faces, the predecessor and successor edges when traversing the left face, and the predecessor and successor when traversing the right face. In Fig. 1, if the start and end

vertices of edge a are X and Y , respectively, the predecessor and successor edges when traversing the left face are b and d, respectively; and the predecessor and successor edges when traversing the right face are e and c, respectively

We can show Fig.1's data structure partly with Table 1.

Table 1. Partly Data Structure of Fig.1

Edge	Vetex		Face		Left Tra.		Right Tra.	
	Start	End	Left	Right	Pred.	Succ.	Pred.	Succ.
a	X	Y	1	2	b	d	e	c

3 The Problem in Our Plan

Fig. 2-1 is a tetrahedron of four vertices A,B,C and D, six edges a,b,c,d,e and f, and four faces 1,2,3 and 4. (Face 3 is back face ACD, face 4 is bottom face ABC; Edge f is AC.).

Now, let us suppose Fig.2-1 is initial state of the tetrahedron, then we revolve it, the final states(one of the face should be in horizontal)can be departed to two kinds. One can be shown in Fig.2-2 to Fig.2-6, the other can be shown in Fig.3. For each state of Fig.3 expresses a state set which includes six more detailed states if we replace vertex D with vertex A in Fig.2. In order to describe those different states in Fig.2, we use table 2,table 3, table 4 to express Fig.2-1 and Fig.2-5, Fig.2-2 and Fig.2-6, Fig.2-3 and Fig.2-4 respectively.

Table 2. For Fig.2-1 and Fig.2-5

E.	Vetex		Face		Left Tra.		Right Tra.	
	S.	E.	L.	R.	P.	S.	P.	S.
b	A	B	1	4	c	a	f	d
f	A	C	4	3	d	b	a	e
d	B	C	2	4	e	c	b	f
a	A	D	3	1	e	f	b	c
c	B	D	1	2	a	b	d	e
e	C	D	2	3	c	d	f	a

Table 3. For Fig.2-2 and Fig.2-6

E.	Vetex		Face		Left Tra.		Right Tra.	
	S.	E.	L.	R.	P.	S.	P.	S.
f	C	A	3	4	a	e	d	b
d	C	B	4	2	b	f	e	c
b	A	B	1	4	c	a	f	d
e	C	D	2	3	c	d	f	a
a	A	D	3	1	e	f	b	c
c	B	D	1	2	a	b	d	e

Table 4. For Fig.2-3 and Fig.2-4

E.	Vetex		Face		Left Tra.		Right Tra.	
	S.	E.	L.	R.	P.	S.	P.	S.
d	B	C	2	4	e	c	b	f
b	B	A	4	1	f	d	c	a
f	C	A	3	4	a	e	d	b
c	B	D	1	2	a	b	d	e
e	C	D	2	3	c	d	f	a
a	A	D	3	1	e	f	b	c

Table 5. Data Structure of Tetrahedron

E.	Vetex		Face		Left Tra.		Right Tra.	
	S.	E.	L.	R.	P.	S.	P.	S.
CAV								
CBV								
CCV								
top vertex								

After observing carefully, we found those tales have some perfect properties as bellow:

.In each table, the items in the second column of the third, fourth and sixth rows are the most left vertex(in the horizontal plane, we will call it class A vertex CAV) in the corresponding figures. For example, in Table.2, this item is vertex A, so the most left vertex in the corresponding Fig.2-1 and Fig.2-5 is vertex A.

In each table, the item in the second column of the fifth row is the second left vertex(in the horizontal plane, we will call it class B vertex CBV) in the corresponding figures. For example, in Table.2, this item is vertex B, so the second left vertex in the corresponding Fig.2-1 and Fig.2-5 is vertex B.

In each table, the items in the third column of the sixth row to eighth row are the top vertex of the tetrahedron in the corresponding figures. For example, in Table.2, this item is vertex D, so the top vertex in the corresponding Fig.2-1 and Fig.2-5 is vertex D.

So when we revolving the tetrahedron, we can change the table data structure to a new state in order to express the final state. The changing rules are listed as Table.5 shows. In Table.5, CCV is the last vertex other than CAV,CBV and top vertex of the tetrahedron. After changing, if the start and end vertex of an edge do not change, such as the edge b in Table.2 to Table.3, than we only copy the other items to the destination row of the final table. Otherwise, we should change the left face and the right face, the left traversing and the right traversing respectively as the Table.3 shows about edge f and d.

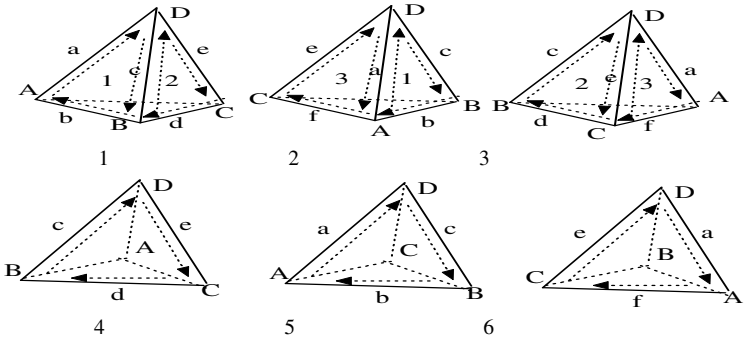


Fig. 2. States of Tetrahedron

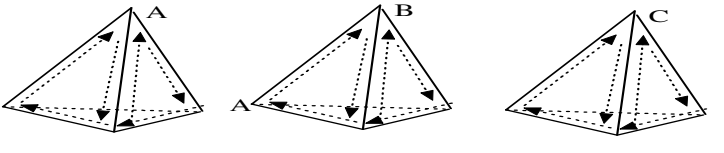


Fig. 3. States of Tetrahedron

4 Design Teach Plan on the Tetrahedron Revolving Problem

Class design. When the coordinates of A,B,C and D are given, we can construct the winged- edge data structure to express the initial state of the tetrahedron.

Problem 1: the constructor of the class should include the function that can check the coordinates of the tetrahedron are legal or not. For example, if the initial state is as it shows in Fig.2-1, then the vertical z-coordinate values of A,B and C are all zero. The constructor function should include some codes to check the condition. The function should throw an exception when an illegal situation happened.

Problem 2: by the coordinates of the tetrahedron, constructor of the class should make decision that which point is CVA, which is CVB and which is CVC, and it use them to fill the winged- edge data structure.

Problem 3: a private function should be set with the class in order to change winged- edge data of an object which originated from the class when the object revolving action happened. In order to simplify the function, the class designer may adopt override concept and design two functions which are name- identical. One deal with the situation that an edge does not change it's direction, and the other deal with the others.

Concepts embed design.

The main goal of the course design is to make students grasping the basic object oriented (OO) concepts of JAVA. So every design group must embed those basic concepts, such as encapsulation, class, object, inheritance, polymorphism, abstraction, interface to their design work.

Above-mentioned three problems are only main design problems. In order to include as many as possible concepts of OO, students must think over and over again and then develop some necessary models to interpret those OO concepts.

Cooperation design.

We often take four students to form a team. A team often includes the following roles.

Team leader: who lead his/her team to finish the course design.

Class designer: who will do overall design of the class, including the properties and methods of the class.

Code writer: who writes the real java code according to the instructions of the class designer.

Code tester: who does the testing work of the whole product.

5 PBL Teach Plan Analyzing

We insist cooperation design is indispensable because much of programming workforce operates in teams, where the habit and ability to work together and to communicate well are a basic breeding for the begin learner of program, especially the OO programming learner. In fact, in order to arrive this aim in our teaching practice, project-based learning method may be a more better choice than PBL. But, even if we take a very simple and small project as practicing material, students have to take a lot of energy to analyze the requirements, to design their problem data structure, to do their total and detailed analysis and so on. If the teach staff ask them to add almost important OO concepts into their design, this may be an impossible question for them to finish within only two weeks.

PBL[4], its literature indicates as if we learn something through a well-established problem solving. In fact, it is broader than this. For example in our plan, we only give a real tetrahedron revolving problem, but other problems, such as within a team, which person take the leading role, the designing role, the coding role is not mentioned; which concepts should be embedded in final work and how embed those concepts into final work is also not mentioned. Those problems are TRUE problems in our plan. All in all, in our plan, information on how to tackle the TRUE problems is not given, although our material are available to assist the students to clarity what the problem consists of and how they might deal with it.

In our experience, we sometimes encounter such situations. One 'GOOD' student in a team does everything of the work. This violates our plan goal. So teacher should do something to avoid it. In our practice, we require class designer submitting class design report, tester submitting test report to avoid this situation.

We found the plan has fowling advantages:

Good documentation training. If every student does a solo copy work like we used to do, they will have less motivation to pay close attention on documentation because every role work in the PBL model should be done by oneself in this non-PBL model. According to our plan, every role member in the team has to pay much attention on documentation, because on one hand students have to note down some his role work in order to let other role member use it, and on the other hand every student should summit an unique documentation report at end.

Integrated curriculum training. In classical studying, we often offer one and another concept to students and so they often have the difficult on integration the different parts of their knowledge. In our plan, we give a single data structure, but this structure can express complex problem in mathematics, so students have 'opportunities' to embed as many as possible OO concepts into their design work. Once this work is done, we achieve the goal.

To work in group training. The ability to work in a group is important for students when they become employers in future. In order to finish individual work in a group, every member must communicate with others. Sometimes one can demonstrate his half-finished work to another, during which each of them must speak. Some possible mistakes maybe found and some new idea maybe produced during the communications.

6 Conclusions

In order to produce a perfect result in our JAVA Course Design lesson, we set an unique course design plan in our practice. Based on students' ability, we take a winged-edge data structure as the infrastructure of the plan. Because this data structure can express a complex problem with a single data structure, so students can use it and only it to embed as many as possible OO concepts into their design work. We also make working in group as a part of our training plan, and adopt the PBL as our main learning method. The experiments told us this is good choice in teaching practice.

References

1. Lowenthal, J.N.: Project-Based Learning and New Venture Creation. In: The NCJIA 10th Annual Meeting, March 23-25 (2006)
2. Falkner, N., Sooriamurthi, R., Michalewicz, Z.: Puzzle-Based Learning for Engineering and Computer Science. Computer 43(4), 20–28
3. Neperud, B., Lowther, J., Shene, C.-K.: Visualizing and Animating the Winged-Edge Data Structure. Computers & Graphics 31(6), 877–886 (2007)
4. Kay, J., Barg, M., Fekete, A., et al.: Problem-Based Learning for Foundation Computer Science Courses. Computer Science Education 10(2), 109–128 (2000)

The Research on CET Automated Essay Scoring Based on Data Mining

Hao Jiang[1], Guoqiang Huang[1], and Jiangang Liu[2]

[1] School of Computer Science and Engineering, Southeast University, 211189, Nanjing, Jiangsu, China
hjiang@seu.edu.cn, huangguoqiang0823@163.com
[2] School of Foreign Languages, Southeast University, 211189 Nanjing, Jiangsu, China
jhonliunj@21cn.net

Abstract. At present, the studies in foreign on natural language processing for automated essay scoring are in full swing. However, these studies are aimed at native English speakers, and it is essentially different from the focus on domestic CET essay scoring. With large-scale popularization of the CET automated essay scoring by the Ministry of Education, the problem on essay scoring is becoming the bottleneck of improving efficiency and large-scale popularization. In this paper, from the perspective of data mining, using classification algorithm KNN-based-association, essay is evaluated in both content and language way. Compared with manual scoring, we analyze their difference.

Keywords: Semantic Analyze Vector Space KNN.

1 Introduction

At present, AES (Automated Essay Scoring) studies in foreign have been becoming a hot issue in natural language processing. And now CET using computer is carried out in various universities by Ministry of Education, and English composition, as a large-scale compulsory language test items, can measure the candidate's mastery of language and comprehensive competence. With large-scale popularization of the CET using computer, the scoring of English writing has been becoming the bottleneck of improving the efficiency of large-scale scoring.

To begin with we will provide a brief background on automated essay scoring, then using classification algorithm KNN-based-association, essay is evaluated in both content and language way. Compared with manual scoring, we analyze their difference from the accuracy and efficiency and conclude our work.

2 Analysis of Automated Essay Scoring

Compared with AES, manual scoring is interfered by various factors, which will directly affect the quality of essay scoring. AES not only ensures the efficiency, fairness and justice, but also provides feedback on composition to the candidates in a timely manner to improve the English learning process.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 100–105, 2011.

The scope of this research lies in both content and language, it contains pretreatment on the composition, the feature processing, language analysis, and makes the corresponding related categories forecast, and computes the results of the weighted score, and implements the AES system ultimately.

2.1 Analysis of Essay Content Scoring

There are three method models to score on the essay content: VSM (Vector Space Model); LSA (Latent Semantic Analyze); concept vector space model based on WordNet semantic dictionary. Modeling of the three methods respectively, compare the pros and cons of various methods by the experimental results.

VSM(Vector Space Model): For each of a document, use a vector to express, turn unstructured text into vector form, and use the mathematical approach to deal with text content. However, VSM has a drawback: it assumes it is independence between the text keywords, in practice, it will produce errors for synonyms and polysemy.

LSA (Latent Semantic Analyze): analyze large sets of text using statistical methods, analyze the potential text semantic structure between words, extract and show the semantics of the words, the latent semantic contains all of the linguistic context information in composition. This model uses a dual-mode factor analysis, based on singular value decomposition [3]. Singular value decomposition is based on the matrix having different entities in the ranks, such as feature item-document matrix, feature items and documents will be expressed as an alternative dimension of a vector space, dot product or cosine value of the vector indicates their similarity. This matrix will be broken down into three specific forms. For example, a $t*d$ feature item-document matrix X, can be broken down into three other matrices:

$$X = T_0 S_0 D_0^{\mathrm{T}} \tag{1}$$

This is called the singular value decomposition of X. T_0 and D_0 are called left singular vector and right singular vector matrix, S_0 contains the diagonal matrix of singular values. By singular value decomposition, LSA can effectively eliminate noise, reduce the impact by the ambiguity and synonymous of words, characterize the true meaning of text better.

Concept vector space model based on WordNet semantic dictionary: WordNet will ignore some of the smaller function words set in English. According to the meaning of the word, form a semantic network, including the synonymous relationship and upper-lower relationship. The model selects the synonym set (Synsets) as a feature item, expresses composition from the semantic aspects. In addition, there also need to use weight function, the product of the concept of frequency and general degree, to enhance the characterization capabilities of feature items.

2.2 Analysis of Essay Language Scoring

For essay language scoring, considering every point of text, such as vocabulary, grammar, sentence structure, select some characteristics variables; conduct a comprehensive analysis of these characteristic variables to measure the quality of the composition language, and score according to each feature item.

Vocabulary. According to the outline CET glossary, give a statistical analysis on new words and the number of the correct spelling of words appeared in the composition, the analysis result could become a feature that can measure the ability of using vocabulary.

Grammar. Select the number of sentences having no syntax errors in the composition as a feature, to measure candidates' mastery of English grammar.

Sentence Structure. Select the number of the different types of phrases in the composition, to measure the candidates' ability of grasping the sentence structure. Select the number of modal verbs to measure the candidates' ability of grasping advanced grammar such as subjunctive.

Apart from the above point of view, we can also choose the number of conjunction and the average sentence length appearing in the composition as a measure of continuity of the sentence. These variables are gathered together, can serve as a standard that can measure the quality of the composition language of candidates.

3 Design and Implementation of AES System

3.1 Generation of Feature Vectors of Essay Content

(1) Do a pretreatment of the composition, it contains isolating from words essay and skipping irrelevant words, such as "a", "the", "is", then turn essay text into the form that computer can handle.

(2) Reduce the dimension of essay: Depending on the feature extraction method (i.e., TFIDF, mutual information, information gain, chi-square test), construct different evaluation functions, conduct an independent assessment for each feature of the initial vector, to obtain an evaluation score. Then according to the size of assessment scores, all the features are sequenced and we select a specified number of feature subset.

(3) Generate feature vector of essay: After extracting feature words, give the weight to the training composition sets and the composition waiting for scoring, generate a feature vector for each composition. Among them, the feature vector has a weight for each feature, to measure the scoring contribution of the feature on the essay content.

a_{ii} is the weight of the *i-characteristic* of *j-chapter* essay, this weight function consists of two parts: ①the local weight function *L(i,j)* , is the weight of feature word *i* in document *j*. It selects the logarithm of document frequency of the feature word *i*:

$$L(i, j) = \log(tf(i, j) + 1) \tag{2}$$

tf(i,j) is the occurrences of the feature word *i* in the *j-chapter* essay.

②the global weight function *G(i)*, is the weight of feature word *i* in the whole essay collection and is the entropy that determined by the distribution of the feature items in all the composition collections:

$$G(i) = H(d \mid i) = -\sum_j p(i, j) \log p(i, j) \tag{3}$$

H(d|i) is the conditional entropy for a given feature word *i. p (i, j)* is the probability of the feature item *i* appearing in the *j-chapter* essay. *W(i,j)* represents the importance of the feature word *i* in the composition:

$$W(i,j) = L(i,j) / G(i) = \log(tf(i,j)+1) / -\sum_{j=1}^{n} p(i,j) \log p(i,j) \tag{4}$$

(4) The compositions are expressed as a vector of semantic space, using LSA and SVD (Singular Value Decomposition) model approach.

3.2 Essay Content Scoring Based on the Distance-Weighted KNN Algorithm

When a essay sample set is unbalanced, scoring error will be relatively large through using the KNN algorithm. Therefore, we select a distance-weighted KNN algorithm, this algorithm weight the contribution of the nearest k neighbors, and according to the relative distance of the query point, assign greater weight to closer neighbors.

Assume that all the compositions are expressed as the point of *n*-dimensional space ξ^n . The arbitrary composition x is expressed as the following feature vector: $< a_1(x), a_2(x), \cdots a_n(x) >$, $a_r(x)$ represents the weight of the *i-characteristic* in composition *x*, thus the distance between two instances x_i and x_j is defined as $d(x_i, x_j)$:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{n} (a_r(x_i) - a_r(x_j))^2} \tag{5}$$

In the KNN learning algorithm, objective function value can be a real value. Function argument is the feature vector of composition content, essay content scoring by teacher is the objective function value here, The return value $f'(x_q)$ is the estimate of $f(x_q)$, it is the average of the nearest k training samples away from *x*.

Training Algorithm:

```
For each training sample;
 (x for the composition of the feature vector, f (x)
     for writing scores);
 Add this sample to list training_examples;
```

Classification Algorithm:

```
Given a query instance to be classified x_q;
    (The feature vector that waiting for scoring);
Select the nearest k samples in the list
    training_examples away from x_q , and expressed by
x_1 K x_k
return f'(x_q) ← Σ_{i=1}^{k} w_i f(x_i) / Σ_{i=1}^{k} w_i   w_i = 1 / d(x_q, x_i)^2
```

When the query point $x_q = x_i$, thus $d(x_q, x_i)^2 = 0$.

In this case, assume $f^{'}(x_q) = f(x_i)$.

3.3 Select Variables for Essay Language Scoring

First of all, separate sentences, and do linguistic analysis. And do syntactic analysis and syntax analysis for each sentence, using open source software--Stanford Parser. First of all, we construct a parser, enter the essay text of all students, parse each sentence, to get a tree structure called grammatical structure. At last, extract language characteristic variables from the tree structure, and make a statistic, put data into KNN algorithm. And score on the essay language.

4 Experiment and Discussions

Based on the "Spoken and Written English Corpus of Chinese students" by Qiufang Wen, Lifei Wang and Maocheng Liang, selected a theme topic for the Western fast food composition. In this thesis, of which 200 were selected as training sample set, 70 as the test set. By doing experiments, do a comparative analysis the pros and cons of the various scoring models compared with the teacher scoring results, the results are given in the Fig.1.

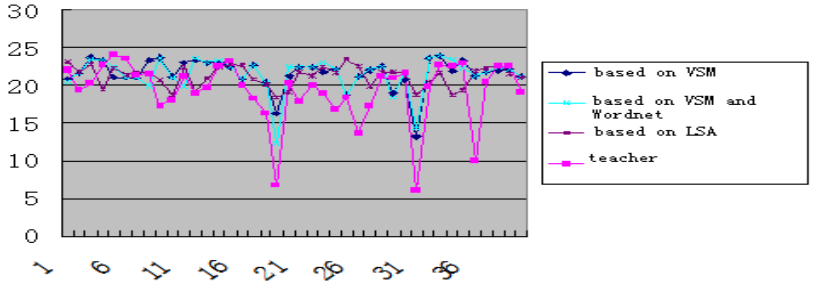


Fig. 1. The results using three models contrast with the teacher scoring

About the errors of the AES with feature selection and teacher scoring, the results are given in Table 1.

Table 1. Comparison of three models

Methods Model	Based on VSM	Based on VSM and WordNet	Based on LSA
Error	411.4	364.4	282.0

Under the condition of a more balanced sample set, we used LSA, selected 30 features using the method of chi-square test, and used the KNN algorithm, to select the minimum error score compared to the essay content waiting for scoring. On whether the feature selection, the error of the scoring results are given in Table 2.

Table 2. The results with/without feature selection

Feature selection	Yes	No
Error under the condition of the average sample set	276.1	343.6

5 Conclusions and Future Work

The results of experiments on English text showed that: in all feature selection methods, Chi-square test and the Information Gain obtained the best results. Term frequency inverse document frequency method has the same performance with the former two, the mutual information method has the worst performance [8-10].

And the result showed that the essay scoring based on LSA was closer to the teacher scoring about essay content, and improved the accuracy of the final score result with the feature selection.

Future work will concentrate more on improvements of KNN algorithm and natural language processing technology for further optimizing the system accuracy and efficiency. Also, we want to improve the deep excavation quality of the composition.

References

1. Dikli, S.: Automated Essay Scoring. Turkish Online Journal of Distance Education 7(1), 49–62 (2006)
2. Landauer, T.K., Laham, D., Foltz, P.W.: Automated Essay Scoring: A Cross Disciplinary Perspective. In: Automated Essay Scoring and Annotation of Essays with the Intelligent Essay Assessor, Mahwah, United States, pp. 87–112 (2003)
3. Singular value decomposition, http://en.wikipedia.org/wiki/Singular_value_decomposition
4. Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D.: Miller,k.: Introduction to WordNet: An on-line Lexical Database. In: Five Papers on WordNet, CSL Report, Cognitive Science Laboratory. Princeton University, Princeton (1993)
5. Rogati, M., Yang, Y.: High-Performing Feature Selection for Text Classification. In: Proceedings of the 11th ACM Conference on Information and Knowledge Management, McLean, America, pp. 659–611 (2002)
6. Yang, Y.: An Evaluation of Statistical Approaches to Text Categorization. Information Retrieval 1(1), 76–88 (1999)
7. Wen, Q., Wang, L., Liang, M.: Spoken and Written English Corpus of Chinese students 2.0. Foreign Language Teaching and Research Press
8. Yang, Y., Pedersen, J.P.: A Comparative Study on Feature Selection in Text Categorization. In: Proceedings of 14th International Conference on Machine Learning, Nashville, United States, pp. 412–420 (1997)
9. Yang, Y., Liu, X.: A Re-examination of Text Categorization Methods. In: Proceeding of the 22nd Annual International ACM Conference on Research and Development in Information Retrieval, Berkeley, United States, pp. 67–68 (1999)
10. Yang, Y.: An Evalutation of Statistical Approaches to Text Categorization. Information Retrieval 1(1), 76–88 (1999)

A Fast Motion Compensation Method in UAV Videos

Zhiyong Ji, Yulong Tian, Jianjiang Lu, and Zhuang Miao

Institute of Command Automation,
PLA University of Science and Technology,
Nanjing 210007, China
Miipl606@gmail.com

Abstract. Unmanned aerial vehicles equipped with surveillance system have begun to play an increasingly important role in recent years, which has provided valuable information for us. In UAV videos, due to the continuous motion of the camera, we can hardly use the traditional background modeling or background-differencing methods to find the moving foreground in videos. Generally, we can utilize feature-based global-motion estimation to achieve this goal. It is unavoidable to balance the time performance and matching accuracy. Although SIFT feature descriptor is invariant to scale and illumination variation, system is difficult to process videos online due to the mass time taken up by extracting SIFT operators. Our thesis quantify the effects of change detection using PSNR and make comparative experiments between the method of using different scales of SIFT pyramids and Harris-SIFT based method. Ultimately, the time performance has been enhanced while maintaining PSNR in a high level.

Keywords: Global-motion estimation; UAV; feature-point detection; change detection estimation.

1 Introduction

In many surveillance applications, people often restrict scenarios, and maintain the posture of cameras in a long time, for example, the surveillance of road condition in traffic management, the surveillance of different areas in malls in order to assemble customers' habits and so on. However, in UAV videos, due to the unavoidable continuous various motions of the camera hung below the aircraft, the background changes frequently so that we can hardly use the traditional background modeling or background-differencing methods to find the moving foreground in videos but resort to seek the relationship between adjacent frames so as to unify correlative images in one reference system.

Global Motion Compensation can solve the problem introduced by camera's motion so as to detect the moving foreground [1]. At the same time, global-motion estimators are an important part of current video-analysis systems such as COCOA system [2] and disaster management [3], in which the most important research topic is about detection, recognition and tracking of moving targets in UAV-captured videos. The methods of Global Motion Compensation can be divided into two categories including global match based motion estimation methods and feature points based

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 106–112, 2011.

global motion estimation methods. The former methods define target function using velocity field by pixels in images and then solve the problem by way of optimizing numerical value [4], which are based upon global matching and complex in computing [5]. Feature-based motion estimators use the motion vectors obtained from a set of selected points to calculate the parameters of the global motion model. This involves the detection of feature points, the computation of correspondences between two sets of features, and the motion parameter estimation. Broszio put forward three steps of feature points based global motion compensation, which are efficient and applied widely [6].

It is obvious that it will reduce time-cost by reducing extracting the number of feature points, however, if there are not enough feature points in each image, the consequent feature matching precision will be affected. So we should keep a balance between precision and efficiency. Farin evaluated the performance of different feature detection algorithms, and concluded that the Harris detector [7] showed the best performance for both repeatability and accuracy [8]. However, Harris detector would not fit for mass rotation changes in UAV videos and it would result in unexpected matching effect at some time. Paper [9] took a compromise way to solve the problem. They utilized the fast extracting feature of Harris detector and the merit of SIFT detector to describe the interesting feature points which obtained good experiment results.

SIFT descriptor is a very robust and reliable representation for the local neighborhood of an image point. In [10], the performance of SIFT descriptor outperforms many other descriptors. Because we need higher matching performance in UAV videos than other situations, so we continue to use SIFT detector [11] to achieve global motion compensation and the RANSAC based robust parameter estimation is examined [12]. For time-critical applications, the scale-space analysis required for the calculation of the SIFT features are too slow (500 – 600ms for images of size 640×480 on a 3GHz CPU) [9]. In order to overcome the abundant time cost by extracting SIFT detectors, we scaled the adjacent images before extracting SIFT feature points and greatly improved the time efficiency.

2 Featured-Based Motion Compensation

Video sequences usually comprise a mixture of foreground object motion and camera motion. Many video-object segmentation algorithms also rely on camera-motion compensated input, because they assume a virtually static camera. The motion field of a rotational camera can be modeled with a projective transform [8]. In this paper, we mainly consider feature-based motion estimation methods, which use the motion vectors obtained for a set of selected points to calculate the parameters of the global motion model. This involves the detection of feature points, the computation of correspondences between two sets of features.

The process of feature point detection identifies a set of feature points in each of the input images. Feature points are placed at locations with unique image content that can be easily found back in the successive frame [8, 11].

The second step establishes correspondences between the feature points detected in a pair of images. Each correspondence can be considered as the motion vector of this

specific feature point. Because the feature points were selected to show unique image content, these correspondences have a high probability to show the correct motion.

The third step estimates the motion-model parameters from this set of feature point correspondences. The main problem in this third algorithm is to separate vectors belonging to foreground motion from background motion vectors. This problem is approached with a robust estimation algorithm that identifies the largest subset of motion vectors that can be described with a single set of camera motion parameters. Under the assumption that camera motion is the dominant motion in the image, the robust estimation will extract the correct camera-motion parameters. This separation is often based on the RANSAC algorithm [12], which is a probabilistic algorithm that only succeeds with a certain probability.

After separation using RANSAC algorithm, the next step is to estimate a full homography matrix with a least squares approach. In an iterative procedure, an affine transformation is computed, filtering outliers in each iteration. And in the final step a full homography is estimated.

In addition, the computation of the nearest neighbor for the purpose of feature matching is very time-consuming. To speedup the nearest neighbor computation, a kd-tree is used to partition the search space [9]. One kd-tree is built for each image. In order to perform the search efficiently, the Best Bin First strategy is used. This algorithm only visits a fixed number of n leaves. The result is either the actual nearest neighbor, or a data point close to it.

3 An Improved Methods by Scaling Images

Although the method of feature-based global motion compensation is to find enough matching points in correlative images and then estimate the parameter of affine matrix, it will spend extra time if system maker lack proper design. Lowe had designed the parameters in practice and we make use of this group of parameter as base to do comparison. Our experiment takes the video clip named actions1.mpg in UCF dataset [13] whose resolution is 960×540. We take the first 500 frames to extract their SIFT feature points, the average feature points in each frame is 926 and the average time consumed in extracting feature points and matching is 1518ms which can hardly satisfy the demand for time efficiency.

3.1 Compensation Effect Estimation by PSNR

The quality of global motion compensation can be evaluated by the effects of change detection, which is the result of subtracting between reference frame and current frame, as illustrated in Fig. 1, in which the results are coming from UCF datasets [13] and have not been done further threshold processing. As a result of illumination changing, the quality of ego-compensation and the differing moving of objects, the effects of change detection are different. From people's vision, the detection quality in frame 41, 55 and 137 is well, in which the foreground can be extracted easily through a proper threshold. In frame 45 and 82, though the quality of motion compensation is also well, the motion of foreground objects is unobvious so we can hardly find them, however, some proper lower threshold can be selected to segment

the moving objects from background. In frame 80 and 104, as the reason of bad motion compensation, a plenty of noises occurred, the grayvalue of these noises are similar to the moving objects, so we cannot take the advantage of threshold method to separate them. In addition, in frame 108 and 109, severe error occurred and system couldn't accept the situation like this. Consequently, we need an objective method to estimate the quality of global motion compensation.

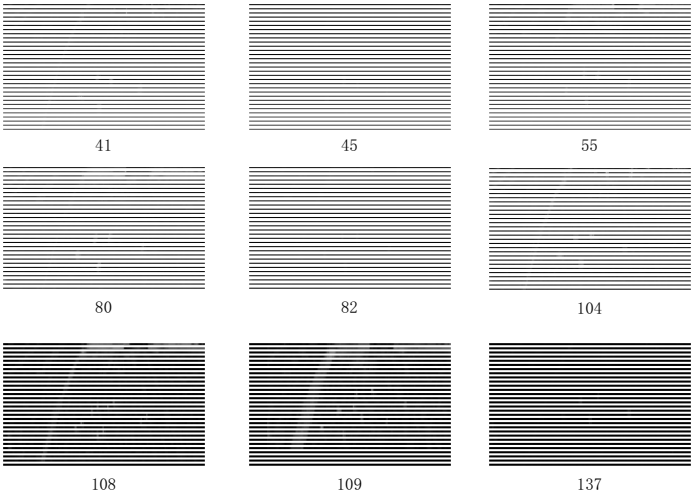


Fig. 1. Examples of change detection effects

Generally speaking, we can separate the estimation into objective methods and subjective methods. Subjective methods call for some observers, who judge the quality of results of the change detection and give out scores according to their experience or some given estimation rule. However, this kind of method requires substantial workload which is fussy and uncertain, so it is constrained in applications. Whereas, the objective quality metrics resort to mathematic modeling to estimate the experiment results which are faster and more stable.

As the most popular quality metric for video quality assessment, we can define peak signal-to-noise ratio (PSNR) to weigh the quality of test images through the deviation from testing image to normalized image in change detection assessment [14]. Given out a digital image $f_0(x, y)$ of the size of $M \times N$ and a reference image $f(x, y)$ of the same size, the value of PSNR is defined as follows,

$$PSNR = 10 \lg \frac{(2^n - 1)^2}{MSE} \tag{1}$$

Where $(2^n - 1)^2$ stands for the square of the most likelihood of signal value, n stands for the bit number of each pixel. In gray images, n=8, $(2^n - 1)^2$=65025.

Many of the objective metrics typically produce a normalized quality estimate ranging from 0 to 1. In contrast, PSNR is computed from MSE using Equation 2, and it can give values up to infinity in case the sequences are identical(MSE=0).

Generally speaking, the bigger the value of PSNR is, the higher the image's quality is. MSE stands for the mean error between observation and expectation as follows,

$$MSE = \frac{\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}[f(x,y) - f_0(x,y)]^2}{MN} \tag{2}$$

Having a mathematic model to measure the quality of the result of change detection, we also have to find a stand test set to finish our experiment. The expected result is too idealistic while the handwork is depleting. For estimating variety of methods in one system, our standard test set is made as follows. First of all, we take the first 500 frames out of one of the video clips in UCF dataset to carry out change detection. The first row in fig. 2 shows some detection results named the experimental observation of PSNR. Afterwards, we choose 30 as the threshold to separate the images into foreground and background, and get the expectation value of PSNR illustrated in the second row of fig. 2, which are also named the experimental expectation of PSNR.

Through experimentation we get the standard average PSNR value of the 500 frames, which is 31.95. Our proposed method is to find a proper way to improve the time performance while keep the PSNR in a high level.

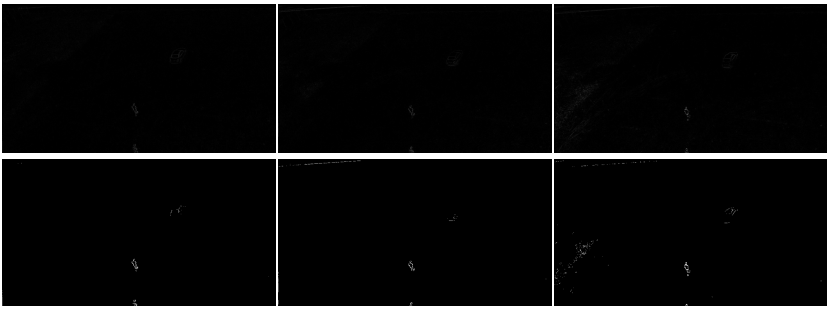


Fig. 2. Illustration of PSNR experimental observation and expectation results

3.2 Experiments and Conclusion

By the leading of reducing the size of image can improve the time performance, the paper had done some comparable experiment by changing the bottom size of SIFT image pyramid and compared it with the method of Harris-SIFT, acquired a little better performance. On the other hand, the more the matching pairs are, the more the system time will waste, at the same time, too few matching pairs will also reduce the compensation precision. So it is our emphasis to balance the number of matching pairs and stabilize the level of PSNR in this section.

Through experimentation, we find the layer of pyramid is the most important key in affecting the time performance in extracting feature points. For the video clip action1.mpg in UCF dataset, the size of basic frame is 960×540, if we build a Gaussian pyramid using bi-sampling, and assure the side of image in the bottom of the pyramid is no less than 8, then the height of the pyramid is 7 in total. In addition,

Table 1. The effects of different scale to image match between SIFT and Harris-SIFT

Method-Pyramid Layers	Size of Pyramid's bottom image	Average Number of extracted feature points	Average Number of inner points	Average time cost of change detection	PSNR
SIFT-8	1920×1080	910.0	546.0	1798.82ms	31.95
SIFT-7	960×540	203.0	139.0	394.79ms	31.81
SIFT-6	480×270	59.0	43.0	124.01ms	31.61
Harris-SIFT-1	480×270	131.0	35.0	124.24ms	25.98
SIFT-5	360×200	31.0	19.0	84.61ms	31.04
Harris-SIFT-1	360×200	106.0	32.0	100.42ms	25.58
SIFT-5	240×135	14.0	7.0	33.09ms	27.93

we can double the size of the bottom of the pyramid to get more feature point and based on which to make comparison. In this section, we still take the first 500 frames mentioned above to make comparison. The average time cost of change detection includes the time cost of feature extracting and matching, frame differencing and denoising, in which the time of feature extracting and matching takes up 98.1% time in total.

We find out that when the height of image pyramid is decreasing to 5, the size of image in the bottom of the pyramid is 240×135 and some mismatching images appeared. When the height of the pyramid equals 6, all images match well. So we can consider the optimal solution is between 240×135 and 480×270 to some extent.

The number of low-quality detection results takes up 22.6% in total when the resolution of image is 240×135 and the effect is too unacceptable, so we take the resolution of 360×200 to make preference, and make the SIFT-method in comparison with Harris-SIFT, from which we can find out that the proportion of inner feature point using Harris-SIFT is far less than that of using SIFT, the PSNR is lower while the time cost is a little more.

Another fact we cannot neglect is that when the resolution of the bottom of the pyramid changes, the affine matrix parameter get from the scaled images cannot be applied in original image pairs to finish change detection. Luckily, there are some potential relationship between the scaled parameter and original parameter. Take affine transformation for example, assuming the pixel coordinate in original image F_1 is denoted as $\{(x,y) \mid (x,y) \in F_1\}$, after affine transformation, the revised coordinate in image F_2 is denoted as $\{(\hat{x},\hat{y}) \mid (\hat{x},\hat{y}) \in F_2\}$. After resizing image F_1, the pixels in $F_1^{'}$ can be denoted as $\{(x^{'},y^{'}) \mid (x^{'},y^{'}) \in F_1^{'}\}$ and correspondingly, the pixels in scaled revised image can be denoted as $\{(\hat{x}^{'},\hat{y}^{'}) \mid (\hat{x}^{'},\hat{y}^{'}) \in F_2^{'}\}$. Meanwhile, the pixels above suffice such conditions: $x^{'}scale_1 = x$, $y^{'}scale_2 = y$, $\hat{x}^{'}scale_1 = \hat{x}$ and $\hat{y}^{'}scale_2 = \hat{y}$. The pixels in image $F_1^{'}$ and $F_2^{'}$ satisfy equation 3,

$$\begin{pmatrix} \hat{x}^{'} \\ \hat{y}^{'} \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^{'} \\ y^{'} \\ 1 \end{pmatrix} \tag{3}$$

which can be transformed as

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ 1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 \times scale_1 \\ a_4 & a_5 & a_6 \times scale_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{4}$$

Lastly and finally, even if the parameters are set properly, there will exist some unavoidable mismatching images and cause severe distortion . In the future work, we will continue to discuss how to take some strategy to detect and repair such errors.

References

1. Keller, Y., Averbuch, A.: Fast gradient methods based on global motion estimation for video compression. IEEE Transactions on Circuits and Systems for Video Technology 13(4), 300–309 (2003)
2. Ali, S., Shab, M.: COCOA - Tracking in aerial imagery. In: Proceedings of SPIE Airborne Intelligence, Surveillance, Reconnaissance(ISR) Systems and Applications, Orlando (2006)
3. Quaritsch, M.: Networked UAVs as aerial sensor network for disaster management applications. Elektrotechnik & Informationstechnik(e&i) 127(3), 56–63 (2010)
4. Su, Y., Sun, M., Hsu, V.: Global motion estimation from coarsely sampled motion vector field and the applications. IEEE Transactions on Circuits and Systems for Video Technology 15(2), 232–242 (2005)
5. Smolic, A., Hoeynck, M., Ohm, J.R.: Low-complexity global motion estimation from P-frame motion vectors for MPEG-7 applications. In: IEEE International Conference on Image Processing, vol. 2, pp. 271–274 (2000)
6. Broszio, H., Grau, O.: Robust estimation of camera parameters pan, tilt and zoom for integration of virtual objects into video sequences. In: International Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging, Santorini, Greece, pp. 15–17 (1999)
7. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of Alvey Vision Conference, pp. 147–151 (1988)
8. Farin, D., Peter, H.N.: Evaluation of a feature-based global-motion estimation system. Visual Communications and Image Processing 5960, 1331–1342 (2005)
9. Azad, P., Asfour, T., Dillmann, R.: Combining harris interest points and the SIFT descriptor for fast scale-invariant object recognition. In: Intelligent Robots and Systems, IEEE/RSJ International Conference on Digital Object Identifier, pp. 4275–4280 (2009)
10. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, pp. 257–263 (2003)
11. Lowe, D.: Distinctive image features from scale-invariant key points. International Journal of Computer Vision 60(2), 91–110 (2004)
12. 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 Association for Computing Machinery 24(6), 381–395 (1981)
13. Aerial Action Dataset, `http://server.cs.ucf.edu/~vision/aerial/index.html`
14. Hyunh-Thu, Q., Ghanbari, M.: Scope of validity of PSNR in Image/Video quality assessment. Electronics Letters 44(13), 800–801 (2008)

An Objects Detection Framework in UAV Videos

Jianjiang Lu, Pengfei Fang, and Yulong Tian

Institute of Command Automation,
PLA University of Science and Technology,
Nanjing 210007, China
miipl501@gmail.com

Abstract. Unmanned aerial vehicles equipped with surveillance system have begun to play an increasingly important role in recent years, which has provided a wealth of valuable information for national security and defense system. The automatic understanding technology based on video contents becomes especially important when facing so abundant information. According to the characteristics of UAV videos that moving objects often appear small and background is complex, our thesis makes research among image normalization, histogram equalization, thresholding methods, morphological processing, motion history image and motion segmentation to find out their different effects in foreground detection. What's more, we have designed basic detection method and enhanced detection method in motion objects detection module, which effectively integrates the traditional single-frame detection technology and multi-frame detection technology into our framework.

Keywords: Moving objects detection; UAV; Single-frame detection; Multi-frame detection.

1 Introduction

Objects detection can be divided into static objects detection and moving objects detection, in which static objects detection is achieved by image segmentation and color, grayvalue, edge and texture-based clustering [1] and moving objects detection is achieved by making use of the continuity of moving information to partition areas in images and associate similar moving areas into target objects[2]. Moving objects detection can be divided into three categories including optic flow methods, background differing methods and frame differing methods.

The methods of optic flow need calculating the field of optic flow first which is sensitive to noise and cannot get a precise result, meanwhile, it is not proper to detect real-time moving objects [3]. The methods of background differing find the moving foreground objects by subtracting current frame and background image, and this method can well extract the shape of objects if static background model provided [4]. However, it costs much to rebuild background model for each frame due to the constant moving of cameras. The method of frame differing is the most direct way to detect moving objects, it is widely accepted by many researchers because of its simpleness and efficiency [2, 5], however, there will be some holes in moving objects or the edge of detected entity may be incontinuous around their edges at some time.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 113–119, 2011.

The method of frame-differing can be categorized into single-frame detection technology and multi-frame detection technology according to the number of frames used. Single-frame detection technology detect objects rely on analyzing the difference between objects and background in images mainly achieved by finding a proper threshold to process image after it was filtered. This kind of method often contains morphologic based method, pixel analysis based method, and wavelet based method and so on. When the noise is too much in comparison with the useful information, multi-frame detection technology is needed which mainly utilize the moving objects' continuity and the randomicity of noises to weed out the ostensible targets. In addition, recognition based objects detection technology is coming out in recent years [6, 7].

There are some differences between objects detection in UAV videos based applications and common monitor applications. Most obviously, the moving objects are often small and the targets' intensity is faintness in UAV videos. Nevertheless, researchers have proposed many methods to overcome the difficulty. In paper [8], Saad detect the regions of interest moving objects by extracting object contours and their system is capable of tracking targets as small as 100 pixels. And in paper [9], Andrew obtained their best results by using a combination of intensity thresholding, motion compensation, interest-point detection, correspondence and pattern classification.

Our work is illuminated by above ideas and designed a small objects detection module which contains basic detection module and enhanced detection module in which image normalization, histogram equalization, thresholding methods and morphologic methods contained in based basic detection module and moving history image with motion segmentation contained in the enhanced detection module.

2 Small Moving Objects Detection in Our Framework

The objects detection flow chart we proposed is showed in figure 1 including basic detection module and enhanced detection module. When the shooting environment is not very complex or wicked, morphologic based basic detection module can well exclude the effect caused by noises and find foreground. However, when the illumination changes frequently, cameras often quiver, targets are tiny or objects are moving in a noisy foreground, the contour of moving objects may rupture and blur at some time, in the case, morphologic based methods will not be enough.

In order to overcome the difficulty, we proposed an enhanced detection module that is constituted by motion templates including motion history image and motion segmentation in which motion history image can cumulate the faint changes in each frame and motion segmentation can find the consistent moving foreground [10]. Our basic detection module mainly contains such steps including change detection, image processing and region refining. In our work, change detection gets a differencing image after motion compensation. Image processing gets a foreground mask by extracting objects' contour. Region refining refines objects' area contour through three frame differing.

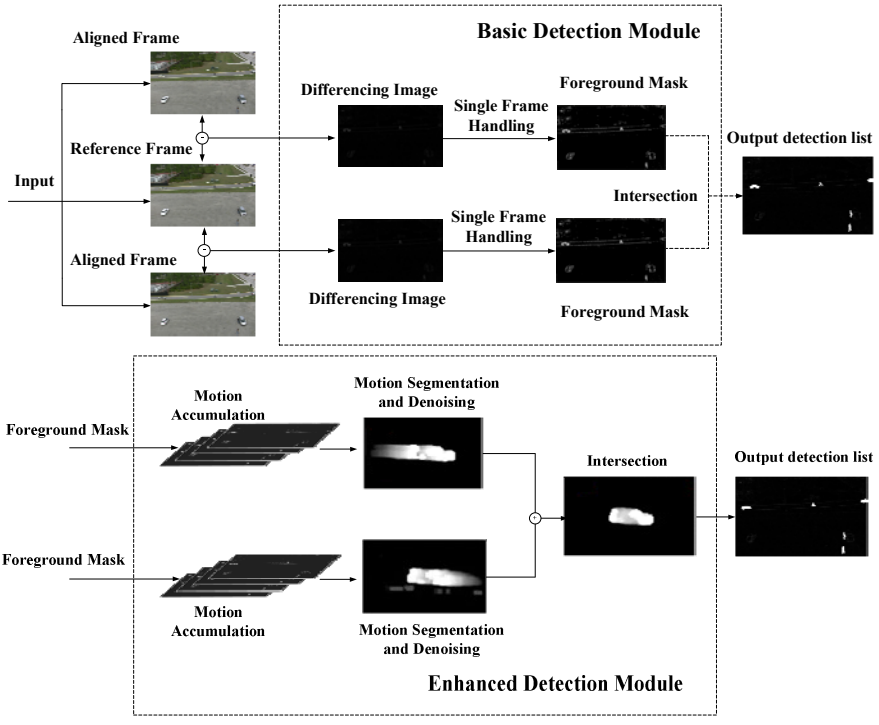


Fig. 1. The framework of our objects detection algorithm

The basic detection module can independently detect the foreground while enhanced detection module can father improve detection capability. When the enhance detection module is needed, the system will not instantly do region refining but process the foreground masks as follow operations including motion accumulation by image history image, motion segmentation to exclude noises and then do region refining at last.

2.1 Change Detection

Change detection finds the difference between two adjacent images by subtracting two aligned images. The input of change detection is current image and reference image, then we align the reference to current image by motion compensation. The output of change detection is a grayvalue matrix constituted by difference gray values from which we can threshold out the moving objects. However, this kind of method needs parallel moving between cameras and targets, and the grayvalue of background keep a constant level in a short time. There are three main approaches in pixel level change detection: background subtraction, inter-frame difference and three-frame difference. However, in the application of UAV video, due to the objects move slowly and the energy of moving foreground is faint, so the method of three-frame difference is not always suitable because it may further lessen the energy of moving foreground.

So we consider if large objects appearing or the velocity of moving foreground is considerable, the effect of three-frame difference is better. Luckily, the forward-backward motion history images help us overcome the problem to some extent because faint energy in each change detection result is accumulated into MHI and it also makes use of the idea of three-frame difference to refine the detection result [5].

2.2 Image Processing

After getting the result matrix by change detection, we need do further processing to find the obvious changing regions and mark them to moving foreground masks. In practice, the masks may contain the moving objects, some static objects and noises which the latter are not needed. A variety of image processing methods are available for us to solve this problem including histogram equalization, image sharpening, normalization, medium/Gaussian filtering, morphological methods and so on. These methods are often applied in different situations.

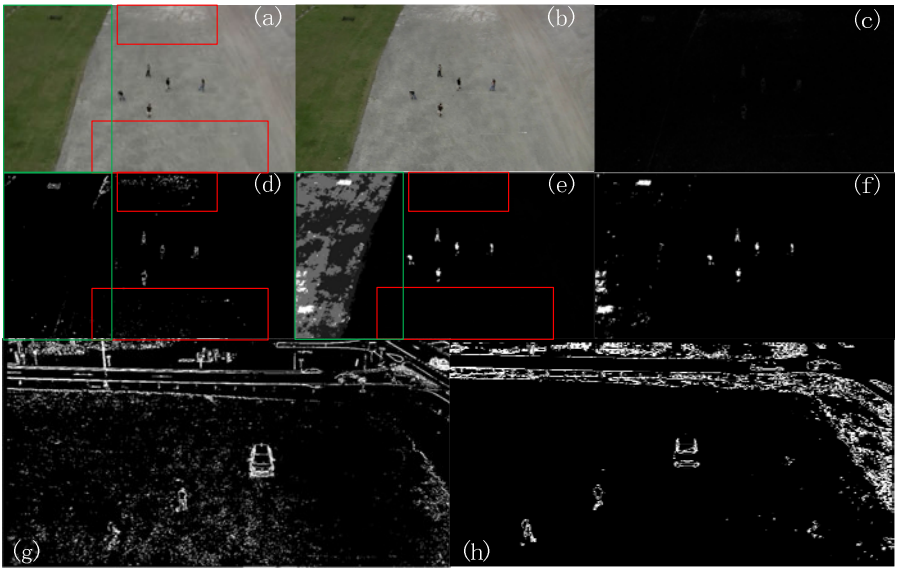


Fig. 2. The effect of normalization to change detection

A notorious problem in airborne video is rapid change in pixel intensities when the camera sensor has automatic gain change control. The gray value of each pixel changes greatly as camera rapidly adjusts its gain to avoid saturation. The changing illumination makes the intensity-based frame difference method inadequate for obtaining accurate motion [5]. So a normalization method is needed and we also take a simple normalization method as forum 1 showed:

$$I'(x,y) = \frac{I(x,y) - \overline{I}(x,y)}{std(I(x,y))} \tag{1}$$

A group of experiments have been done as illustrated in figure 2 in which (a) and (b) represent two adjacent frame, (c) represents the differing result, (d) is the result after differing by the threshold value of 15 in which some pepper-salt noise exist, the method of normalization can weep out this kind of noise but imports some other new noise showed in (e). Figure (g) and (h) is another group of experiment results, (g) stands for not doing normalization for single image before frame-differing and (h) stands for doing normalization for single image before frame-differing. So it is surely the method of normalization can move the pepper-salt noises introduced by illumination. However, when the grayvalue of moving foreground is not obviously different from its around background, the normalization to change detection result may destroy the contour of moving foreground as illustrated in figure 3(d) in comparison with figure 3(c). So it is better to normalize the image by region but not normalize the image as a whole.

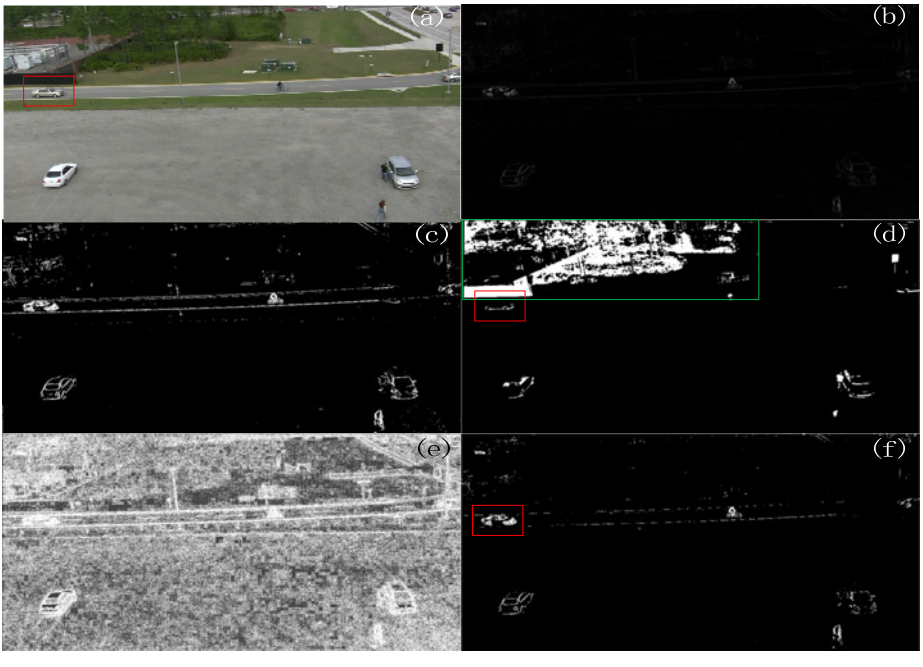


Fig. 3. The effect of white equalization to change detection

Another similar processing method is white equalization also named histogram equalization which is used to make the grayvalue histogram more smoothable. In this paper, we make such steps to achieve white equalization. (1) Count the frequency of each grayvalue appeared in each pixel and calculate:

$$p(i) = \sum_{x=1,\, y=1}^{x=m,\, y=n} \delta[src(x, y) - i] / (m \times n) \tag{2}$$

in which the variable i ranges from 0 to 255. (2)Sum $p(i)$ to form $g(i)$, and $g(i)$ stands for a probability value that grayvalue is less than i and $g(255)=1$. (3)Adjust every pixel by the forum $dst(x,y)=g[src(x,y)]\times 255$ to form the final change detection result. In comparison with the method of normalization, histogram equalization avoid the probability that a large range of noises appearing and can lessen noises at the same time. Other image processing methods about thresholding and morphological methods can reference paper [10, 11].

When the signal to noise rate is very low and random strong noises produce many ostensible targets, single frame detection technology is hard to apply and need multi-frame detection technology to locate moving objects. It mainly includes motion history image and motion segmentation. Motion history image can accumulate faint changes in each differing result [5] and motion segmentation can find consensual moving foreground [10], our enhanced module combine them to overcome the effect of noises and fit for detecting objects of different sizes and velocity.

2.3 Motion History Image

For small objects, because they lack of obvious shape, size and texture information, it's hard to detect moving targets only knowing the foreground mask. On the other hand, the foreground objects are scattered by single frame differing. Motion accumulation [8] or motion history image can be used to get the objects' moving information. In comparison with motion accumulation, MHI not only enhances the foreground, but also keep the information of newest position, and can utilize MHI to do further motion segmentation.

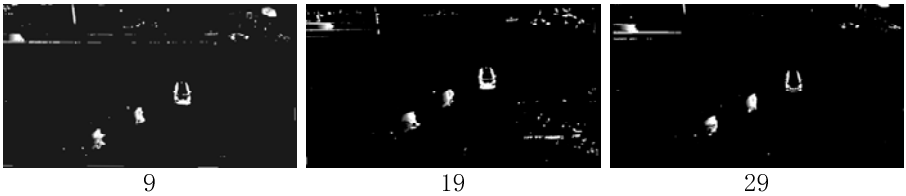


Fig. 4. The detection result using MHI

Motion energy image need to be built before constructing a MHI, Bobick had proposed how to calculate a MEI in paper [12]. However, motion energy image can only present the location of moving objects but cannot describe their moving details. So MHI is also needed. MHI can describe many motion properties such as motion direction. Figure 4 shows the experimental result using MHI to detect moving foreground.

2.4 Motion Segmentation

Motion segmentation is an effective way to track general movement [10]. Although MHI can enhance the faint foreground, it cannot move the effect of noises entirely so we utilize motion segmentation to do further analysis. Motion segmentation aims at

calculating every moving foreground blob's global motion vector in its corresponding MHI's lasting time and it relies on motion gradient and stamp information in MHI. Sequentially fading silhouettes record the history of previous movement and thus are referred to as the motion history image. Once the motion template has a collection of object silhouettes overlaid in time, we can derive an indication of overall motion by taking the gradient of the MHI and we can take these gradients by using the Scharr or Sobel gradient functions. Finally, we can collect a measure of global motion.

The longer time the objects last, the more coherent the objects' motion have, the more obvious the interesting moving targets are. At the same time, the noises appear with no regulation and the absolute value of global motion vector in noises is obviously less than that of extracted in moving foreground. So through statistic the global motion vector in each blob, we can further extract interesting targets.

References

1. Martin, D.R., Fowlkes, C.C., Malik, J.: Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(5), 530–549 (2004)
2. Richard, J., Srinivas, A., Omar, A.: Image change detection algorithms: a systematic survey. IEEE Transactions on Image Processing 14(3), 294–307 (2005)
3. Sun, Z.H., Zhu, S.A., Zhang, D.W.: Real-Time and automatic segmentation technique for multiple moving objects in video sequence. In: IEEE International Conference on Control and Automation, pp. 825–829 (2007)
4. Stauffer, C., Grimson, W.: Learning patterns of activity using real-time tracking. IEEE Transactions Pattern Analysis and Machine Intelligence 22(8), 747–757 (2000)
5. Yin, Z.Z., Collins, R.: Moving object localization in thermal imagery by forward-backward MHI. In: Proceedings of IEEE Computer Science Conference on Computer Vision and Pattern Recognition, pp. 133–140 (2006)
6. Bai, X., Wang, X., Latecki, L.: Active Skeleton for Non-rigid Object Detection. In: International Conference on Computer Vision (2009)
7. Maji, S., Malik, J.: Object Detection using a Max-Margin Hough Transform. In: Proceedings of IEEE Comference on Computer Vision and Pattern Recognition, pp. 1038–1045 (2009)
8. Ali, S., Shab, M.: COCOA - Tracking in aerial imagery. In: Proceedings of SPIE Airborne Intelligence, Surveillance, Reconnaissance(ISR) Systems and Applications, Orlando (2006)
9. Miller, A., Babenko, P., Hu, M.: Person tracking in UAV video. In: Proceedings of International Workshop on Multimodal Technologies for Perception of Humans, Berlin, Heidelberg, pp. 215–220 (2008)
10. Bradski, G., Kaebler, A.: Learning OpenCV. O' Reilly Media, Inc., Sebastopol (2009)
11. Otsu, N.: A threshold selection method from gray-level histogram. IEEE Transactions on Systems Man and Cybernetics 9, 747–757 (1979)
12. Bobick, A.F., Davis, J.W.: The recognition of human movement using temporal templates. IEEE Transactions on Pattern Analysis and Machine Intelligence 3(3), 257–267 (2001)

A Evaluation Method Based on Fuzzy Semantic Differential for Web-Based Customization System

Xiumei Zhang[1] and Yongjian Sun[2,*]

[1] Department of Economics and Trade
Hangzhou Wanxiang Polytechnic, Hangzhou, Zhejiang 310023, China
laura74@163.com
[2] School of Mechanical Engineering & Automation
Zhejiang Sci-Tech University
Hangzhou, Zhejiang 310018, China
yjsun@zstu.edu.cn

Abstract. Consumer's personalized requirements, or the so-called Kansei needs, have become one of the most important concerns in product customization system. Due to the vagueness and uncertainty of consumers' requirements, the Fuzzy Semantic Differential (FSD) method has been used in the process of the customized product design to build the relationships between the consumer's personalized requirements and the feature parameters of the product. The FSD method is used to evaluate the 3-D product model. According to the result of the Kansei evaluation, the feature parameters of the 3-D product are adjusted to suit the customer's preference. This method is used and illustrated using a case study on men's suit. The results show that the proposed method is promising in dealing with evaluating the customer's requirement for the customized product.

Keywords: Product customizaiton system; personalized requirement; Kansei engineering; Fuzzy semantic differential method; Inactive design.

1 Introduction

In recent years, the product customziation system (PCS), as a integrated system combining with the customer's personal requirement, has drawn some researcher's attention [1-3]. The rise of PCS mainly results from the change of the people's consumption concept from purchasing the product to acquiring the utility. People pay more attention to the capabilities of a product or service to suit their personalized requirements. Under these circumstances, how to evaluating the customer's perceiving utility is very important to the providers of the product.

Kansei Engineering(KS), as a methodology to deal with the consumer's subjective impressions(called Kansei in Japanese), mainly used in the process of human-centered design, and quantify the customer's feeling for the design by using some methods, for example, market investigational method, analytic hierarchy process(AHP), Delphi method, comprehensive evaluation method, and semantic

* Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 120–127, 2011.

differential method (SD),etc [4-7]. The most commonly used method is the SD method, among Kansei engineering, Kansei evaluation is an important process in which a product may be selected for production or design [8-9].

In this paper, based on the characteristics and application of the PCS, we establish an evaluation method and process for evaluating the customers' satisfaction for the product and service of the providers , and apply the method to evaluating the customers' satisfaction in the online personalization and customization of the suit. Firstly, we introduce a fuzzy semantic differential method by combining each advantage of the fuzzy theory and the SD method in evaluating the customer's perceiving utility, Secondly, we describe the procedure of using the FSD method in the Web-based customization system. Finally, we apply the method to the customer's evaluation for the 3-D personalized suit through the online customization system.

2 Fuzzy Semantic Differential Method

2.1 Interaction Design Based on 3D Product Model

The interaction process among customer, product and designer play an important role to obtaining the customer requirement as shown in Fig.1. Especially in personalization customization, the customer exchanges their idea with the designer through using a 3-D interactive virtual reality model. Then the designer alters the virtual 3-D model until the customer is satisfied with it. By the evaluation methods including the Fuzzy SD method, the customers' expectation for function or emotion of the product is developed.

2.2 Fuzzy Semantic Differenetial Method

Evaluation for the product is the customer's subjective judgment, despite the information of the product is rational or perceptual. The rational information of the product includes the function, material and production technology of the product, but it depends on the customer to evaluate it to meet his requirement or not.

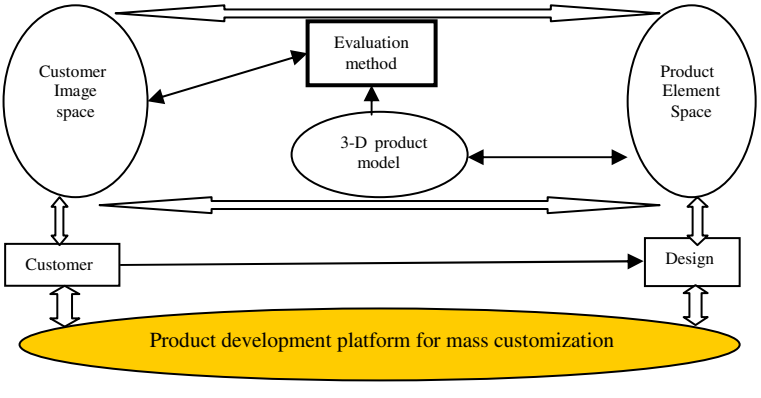


Fig. 1. A framework of Kansei evaluation for the Web-based customization system

Fig. 2. Semantic Differential inventory

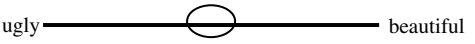


Fig. 3. Fuzzy Semantic Differential inventory

The Semantic Differential (SD) method is created by Osgood and his colleagues [8]. It is designed to measure the connotative meaning of concepts. It is composed of three parts: an object or a concept being evaluated, an adjective and a tester. Fig.2 shows the tester chooses one number in 1-7 to express the degree of beauty of the object. There is a problem that the tester knows the range of evaluation, as shown in Fig.3. In view of this case, we use the fuzzy logic to solve this problem. The "fuzzy logic" is the methodology created in opposition to the classical logic. It is designed to represent the uncertainty and vagueness of the reality. So it can represent the vagueness of the customer's satisfaction. The Fuzzy Semantic Differential (FSD) method is analyzed and explained as show below.

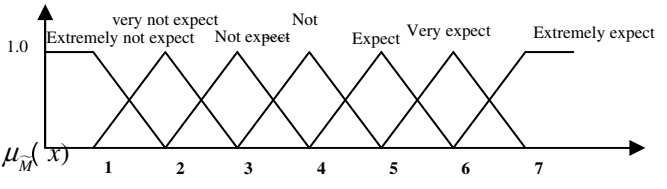


Fig. 4. Fuzzy membership function of the semantic change degree

Step 1. Identification and measurement of Kansei attributes

Let X= $[X_1, X_2, ..., X_N]$ be set of Kansei attributes of products, where N denotes the total number of Kansei attributes, and $K_n = [K_n^{-i}, ..., K_n^i]$ be a linguistic term set for each Kansei attribute of products, when i=3 means a 7-point scale. For example, "expect" is a Kansei attribute of a product having the Kansei words {"extremely not expect",…,"extremely expect"} with a 7-point scale in Fig.2.

$$K_m = [K_m^{-3}, K_m^{-2}, K_m^{-1}, K_m^0, K_m^1, K_m^2, K_m^3] == \{\text{"extremely not expect", "very not expect", "not expect", "not care", "very expect", "extremely expect"}\} \quad (1)$$

In equation (1), m means the m-th attribute of products.

Step 2. Fuzzification of the semantic change degree

The triangular function adopted as the membership function as shown in Fig.4. Fig.4 represents the membership function of the customer's expectation, for

simplicity, the triangle or trapezoidal form of the membership function is used most often for representing the fuzzy number. In this study, triangular fuzzy membership function is used.

$$M_i = (l_i, m_i, n_i) \text{ li} < \text{ mi} < \text{ n i , i=1,..,7} \tag{2}$$

Where M is the triangular fuzzy number of $\mu_{\widetilde{M}}(X_i)$, l, m, n is the real number.

$$\text{M=[} M_1, M_2, M_3, M_4, M_5, M_6, M_7 \text{]=[(1,1,2),(1,2,3),(2,3,4),(3,4,5),(4,5,6),(5,6,7),(6,7,7)]} \tag{3}$$

Where M1, M2, M3, M4, M5, M6, M7, are the triangular fuzzy number of membership function .

$$\mu_{\widetilde{M}}(X_i) = \begin{cases} (X_i - l)/(m-l), l \le X_i \le m \\ (X_i - n)/(m-n), m \le X_i \le n \\ 0 \end{cases} \tag{4}$$

Step 3. Defuzzification

In order to be easy to evaluate the Kansei attribute, it was necessary to compute the Kansei evaluating value by the process of defuzzification. There are many methods, for example, the Centroid method, the Maximizing Set and Minimizing Set method and the maximum degree of membership e method et al.. The Maximizing Set and Minimizing Set method is widely used in relative literatures, and the solution can be figured out quite quickly. This study uses the Maximizing Set and Minimizing Set method which is proposed by Chen to defuzzification [9].

Maximizing Set $R = \{(X_i, \mu_R(X_i)) \mid X \in R\}$, Minimizing Set $L = \{(X_i, \mu_L(X_i)) \mid X \in R\}$

$$\mu_R(X_i) = \begin{cases} (X - X_{\min})/(X_{\max} - X_{\min}), X_{\min} \le X \le X_{\max} \\ 0 \end{cases} \tag{5}$$

$$\mu_R(X_i) = \begin{cases} (X - X_{\min})/(X_{\max} - X_{\min}), X_{\min} \le X \le X_{\max} \\ 0 \end{cases} \tag{6}$$

Right Utility $U_R(\widetilde{M_i})$ and left Utility $U_L(\widetilde{M_i})$ as follows:

$$U_R(\widetilde{M_i}) = \sup(\mu_{\widetilde{M}}(X_i) \wedge \mu_R(X_i)) = \frac{n_i - X_{\min}}{(X_{\max} - X_{\min}) - (l_i - n_i)} \tag{7}$$

$$U_L(\widetilde{M_i}) = \sup(\mu_{\widetilde{M}}(X_i) \wedge \mu_L(X_i)) = \frac{X_{\min} - m_i}{(X_{\max} - X_{\min}) + (l_i - m_i)} \tag{8}$$

Total value of M_i is shown as follows:

$$U_T(\widetilde{M_i}) = (U_R(\widetilde{M_i}) + 1 - U_L(\widetilde{M_i})) = \left(\frac{n_i - X_{\min}}{(X_{\max} - X_{\min}) - (l_i - n_i)} + 1 - \frac{X_{\min} - m_i}{(X_{\max} - X_{\min}) + (l_i - m_i)}\right) \tag{9}$$

3 Procedure of Using the FSD Method in Web-Based Customization System

In order to quickly create the customized 3-D product model to meet the personalized requirement, in this paper, the method based on FSD is used to evaluate the 3-D product model. According to the results of the evaluation, adjustment for the parameters of the template of the 3-D product model is not stop until the customized 3-D product model is acquired. In Fig.5, the representative 3-D product templates are chosen to compose the template library as shown in box ①, through clustering the product features that have the highest Kansei evaluation result. Creating the 3-D product model based on the suitable template is shown in box ②, then the 3-D product model is evaluated by the FSD method as shown in box ④ on the base of identifying and measuring the Kansei attribute of the 3-D product model. According to the result of the kansei evaluation for the 3-D product model, the feature of the product is adjusted to suit the customer's requirement as shown in box ⑤, if the selected template is not suit for the customer's requirement, then the new template is reselected from the template library.

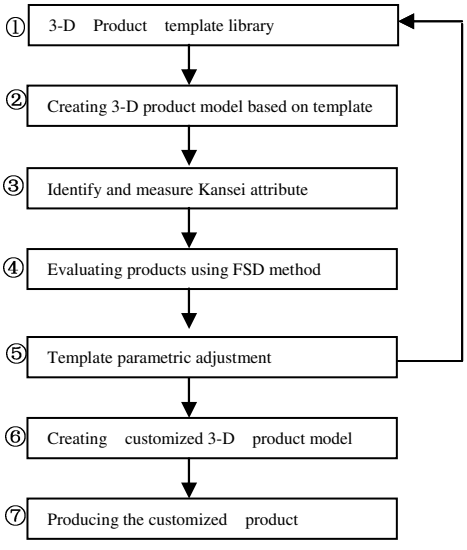


Fig. 5. Procedure of using the FSD method in Web-based customization system

4 AN Application: Kansei Evaluation for Online Customized Suit

This company under study provides the online customized suit for the high-end users in Ningbo, China. 40% of the commitments is from online customization. In the process of selecting their prefer suits, the customers may pay attention to the key

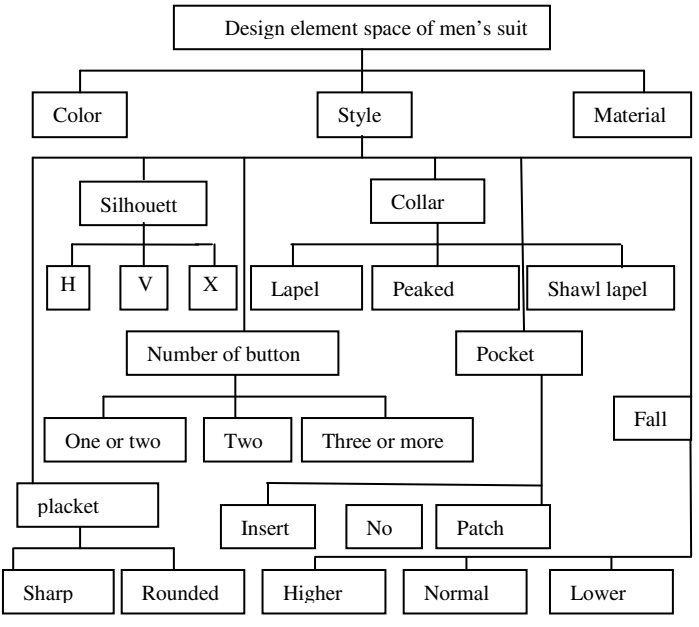


Fig. 6. Framework of the design element space of men's suit

Table 1. Form of the customer's Kansei evaluating for 3-D

Modeling elements		Extre-me	Very	Fairy	Aver-age	fairy	very	Extre-me
Silhouette	H							
	X							
	V							
Collar	lapel							
	peaked lapel							
	shawl lapel							
Pocket	insert							
	no							
	patch							
placket	sharp							
	round							
Fall	higher							
	normal							
	lower							

elements of suit. “color”, “model” and “material”. The Design elements of men’s suit are discussed, analyzed and decomposed by using interviewing and laddering methods and then represented in the form of hierarchy structure so as to build the design elements space of men’s suit as shown in Fig.5. only the design elements of suit style is further subdivided into “Silhouette”, “Collar”, “pocket”, “placket” and “Fall”, “color” and “material” is not further subdivided in this article allowing for the space of article. Before the customization system for suit is developed, a large amount of Kansei adjective words are collected, and through the analysis and filtration, these adjective words are prepared for sorting. Through computing the intimate degree between these words, the semantic space of evaluation for the men’s suit is clustered in the form of 4 word-pairs: formal-casual, classical-modern, gorgeous-simple, and elegant-masculine. The customers evaluate the 3-D suit according to their perceptions, just only use the green circle on the red line as shown in table.1. Then the system will give the evaluation result automatically. By comparing the evaluation result with the reference value set beforehand, then the customers may change the dissatisfactory design element. Then they evaluate the regenerating 3-D suit model, until the customer’s preferred 3-D suit model is obtained.

5 Conclusion

The present work shows that the semantic differential has the disadvantage to use the measure table to evaluate the strong and weak parts of the feeling. The fuzzy logic is suitable to evaluate the uncertainty and the imprecision of the people’s thought, so the fuzzy semantic differential method was called up. It was used to evaluate the customer’s perception and expectation for the personalized product and service on-line, then the degree of the customer’s satisfaction can be obtained. Application of this method to the Internet-based personalized customization system for suits shows that this method has some advantages for evaluating the functional and Kansei requirements. It also would be of important help in the recommender systems of the product and service.

Acknowledgments. We also like to appreciate constructive comments and valuable suggestions from anonymous referees. This research is supported by Zhejiang Natural Science Foundation (R60804030, Z6090572) and the National Natural Science Foundation of China (No.50675208).

References

1. Hirschl, K.W., Scholl, G.: New concepts in product use for sustainable consumption. Journal of Cleaner Production 24, 873–881 (2003)
2. Gebauer, H., Fleisch, E., Friedli, T.: Overcoming the service paradox in manufacturing companies. European Management Journal 23, 14–26 (2008)
3. Lai, H.-H., Chang, Y.-M., Chang, H.-.C.: A robust design approach for enhancing the feeling quality of a product: a car profile case study. International Journal of Industrial Ergonomics 35, 445–460 (2005)
4. Yan, H.-B., Huynh, V.-N., Murai, T.: Kansei evaluation based on prioritized multi-attribute fuzzy target-oriented decision analysis. Information Sciences 178, 4080–4093 (2008)

5. Lai, H.-H., Lin, Y.-C., Yeh, C.-H.: User-oriented design for the optimal combination on product design. International Journal of Product Economics 100, 253–267 (2006)
6. Lin, Y.-C., Lai, H.-H., Yeh, C.-H., et al.: Consumer-oriented product form design based on fuzzy logic: a case study of mobile phones. International Journal of Industrial Ergonomics 37, 531–543 (2007)
7. Llinares, C., Page, A.: Application of product differential semantics to quantify purchaser perceptions in housing assessment, building and environment. International Journal of Product Economics 42, 2488–2497 (2007)
8. O'Hagan, M.: Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic. In: Proceeding of the 22nd Annual IEEE Asilomar Conference on Signals, Systems, Computers, Pacific Grove, Montreal, Canada, pp. 681–689 (2008)
9. Chen, S.H.: Ranking fuzzy numbers with maximizing set and minimizing set. Fuzzy Sets and Systems 17, 113–129 (1985)

The Study of Plagiarism Detection for Program Code

Hao Jiang and Zhemin Jiang

School of Computer Science and Engineering, Southeast University,
211189, Nanjing, Jiangsu, China
hjiang@seu.edu.cn,
yingminxing@126.com

Abstract. With the increasing popularity of programming courses, the cases of plagiarism also rises rapidly as well. Plagiarism detection methods and verifying the originality of student's work program has become particularly important nowadays. By studying similar measurement techniques of existing code, this document focuses on the forward maximum matching algorithm proposed to improve an existing and efficient segmentation method while proposing effective marker string replacement rules in order to shorten the length of the string tag. At the same time, this paper proposes a new marker string generation method – generating tag strings in accordance with each function execution sequence, in order to eliminate redundant functions of the test results. Finally, the system would take the RKR-GST algorithm as a token string matching algorithm. The experimental tests have shown that the improvement over plagiarism detection program code has a significant effect in the long run.

Keywords: plagiarism detection; generating tag strings by the function sequence; RKR-GST algorithm; forward maximum matching algorithm.

1 Introduction

With the rapid development of information technology, attending programming courses offered by the institutions of higher learning will undoubtedly become the preferred method of education, in order to help students obtain the necessary programming skills and practical hands-on skills to excel in the competitive environment in the sector. However, the result of this trend is a large number of students' committing plagiarism in order to complete their assignments tasked to them, which will then increase the workload of the reviewer too. Therefore, developing an automatic detection system for Plagiarism is thus necessary. Upon its creation, it will then be able to complete the necessary tasks quickly and efficiently through the usage of the program code detection system. The procedure code detection system which is described by this paper will convert the program code into tag (token) strings that describe the program characteristics, and then identify similar parts of tag strings through in-depth comparison of them against existing programs created already. Finally, to determine a suitable formula used to calculate the matching results in order to obtain the value of similarity.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 128–133, 2011.

2 Plagiarism Detection of Code

2.1 Attribute Counting Method [1]

Halstead [2] first proposed the idea for the program code plagiarism detection. This is done by defining the similarity between the two blocks by calculating the distance between the Halstead feature vectors of a pair of blocks. The smaller value of distance, the higher the similarity of the codes; thus increases the possibility of plagiarism occurrence.

2.2 Structural Metric

Structural metric method is based on the structure of the program metrics based on the statistical indicators, which determines the degree of similarities between the programs. It is carried out with two steps: firstly, by parsing the program based on the programming language, and then generating a tag string (token) sequence, which converts each of the programs to a string sequence; secondly, by calculating the values of two similar marker sequences according to a matching algorithm. Considering the structural characteristics of the program, some of the "structural level" of the plagiarism detection is better.

3 Introduction of String Matching Algorithm

3.1 Definition of Similarity

Similarity [3] is a quantitative value of two procedures' comparison which is used to determine whether there is plagiarism between two procedures. In general, a similarity is represented by a value (0.0 ~ 1.0) or percentage value (0% ~ 100%). The similarity of code A and B can be measured in the following formula:

$$Sim(A,B) = \frac{\left|S(A,w) \cap S(B,w)\right|}{\left|S(A,w) \cup S(B,w)\right|} \tag{1}$$

S (A, w) and S (B, w) are the sets of all sequences which is in the size of w. The larger of Sim value, the more similar of two codes' composition.

3.2 RKR-GST Algorithm

Based on the GST [4] algorithm, RKR-GST is an avarice string matching algorithm which introduced the KR algorithm. Different from the LCS algorithm [5] and the Levenshtein algorithm [6] and [7], it not only does not require a strict and orderly string, but has a low time complexity too. This algorithm can be used to find out the most common sub-code string of two codes and calculate the similarities of the two sections of code A and B.

4 Improvement of the Marked String Generation Method

4.1 Segmentation Algorithm

Segmentation algorithm is a process which extracts legal vocabulary from the text. Forward maximum matching algorithm intercepts the sub-strings from left to right and gets the longest legal vocabulary by matching with the thesaurus terms. Due to its simplicity and high precision, it is used more often in the sub-word of formal language. There are three factors on the time of the algorithm: the size of lexicon, the values of the maximum string length L, and code specifications.

4.2 Improved Segmentation Algorithm

After analysing the forward maximum matching algorithm, it is found that the unsuitable maximum string length will cause frequent unsuccessful matches, while the size of the lexicon is another important factor on the time of word segmentation. Thesaurus used in this article only needs to retain the language keywords and predefined identifier, so its size will be an effective streamlining one.

This study code to be analyzed has a pre-program process and all the lexical elements in the code are divided into three types: custom identifiers; definition of keywords and identifiers; other characters as well.

Throughout the process, each sub-string compares with the dictionary up to a single time. In the process of the analysis of the whole series, the times to compare with the dictionary will not exceed the number of identifiers through splitting

4.3 Expand Rules of Marked String

Expand by Files

Marked string is composed by tag strings' order of appearance in the file and the tag string corresponds to all function modules. According to the documents, the expansion method generates tag string which has a better plagiarism detection effect on statement-level changes (such as modifying the identifier name), but for function-level changes (such as increasing redundancy function), the detection effect is not satisfactory still at this point of time.

Implementation by the Order of the Function Module

The tag string of code is composed by tag strings' order of implementation and these tag strings corresponds to function modules. The redundant function is a function module which is defined in the code but not used, therefore, according to the order of function module execution; the expansion method can detect and eliminate redundant functions of any corresponding tag strings.

Comparison of the Two Expansions

Without the module decomposition and sorting, expanding files needs line scanning only, so it requires very little storage space and time. But it can-not eliminate redundant

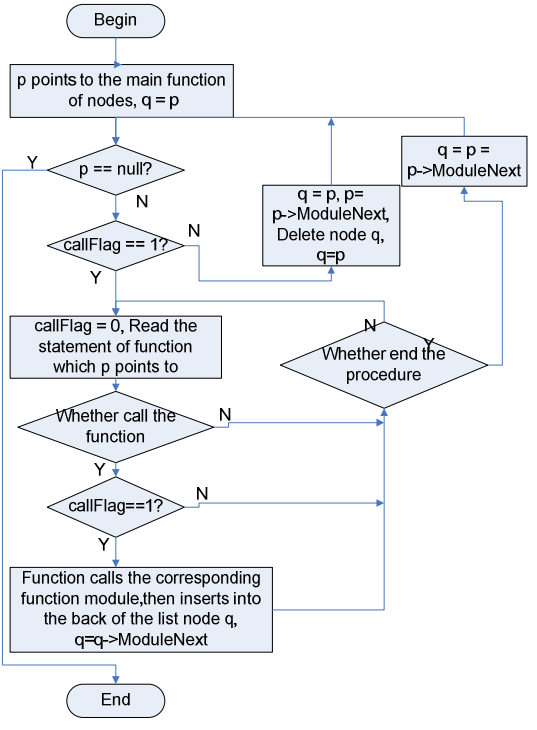


Fig. 1. The flow chart of the sorting process function module

functions, in the case of many redundant functions, the increment of marked string length have a great influence on the test results. Expanding by the order of function execution can eliminate redundant function and reduce the length of marked string. It requires less time and has a better effect on the function level of modification. However, as it requires the module decomposition step and stores the appropriate data type, therefore it needs more storage space and duration for a task to be completed. Considering the fact of the rapid and rampant development of computer hardware and software, the additional cost of storage is not the main controversy happening. Plagiarism detection system focuses on the running time of the formation of a tag string and the matching process. To sum up, the expansion by order of the function module execution is much more desirable than those by file which can reduce the running time of tag strings' comparison algorithm.

5 Experimental Analysis and Comparison

Sample A of the experimental group has some simple changes: only changes in comments, addition of extra spaces and blank lines, and increases surplus statements. Two methods of test execution can achieve the desired effect and time, so the similarity is not comparable. Thus, it would not be described though the comparison results of the following experiment B and experiment C would be studied and observed.

Experimental group B samples were obtained by function-level changes. T6 samples exchange order of the functions, T7 samples add the functions which are not called. The two experimental results obtained in the group of test samples are shown in Table1:

Table 1. Experimental results of the experimental group B

	Execution time		Similarity		
	By function	By file	By function	By file	Expected value
T6	0.5ms	0.54ms	100%	100%	100%
T7	0.53ms	6.8ms	100%	60.7%	100%

Experimental group C samples are obtained by integrated modification. T8 samples use all the integrated methods. The experimental results obtained in the analysis of samples are shown in Table 2:

Table 2. Experimental results of the experimental group C

	Execution time		Similarity		
	By function	By file	By function	By file	Expected value
T8	4.46ms	11.68ms	95.1642%	50.5087%	100%

From the comparison of experiment B and experiment C, it is clear that the sample on the function-level changes, in particular, adding redundant functions, not only requires more execution time as compared to the expansion method by file, but has less degree of similarities as well. For the sample on comprehensive changes, the effect is the same. Therefore, based on the improved segmentation algorithm, the plagiarism detection system which is in accordance with the order of function module execution to generate the marked string has good detection results on the function-level.

6 Conclusion

This paper focuses on the characteristics of the program code. It uses an improved segmentation technology, while a new marked string generation method is introduced------generating tag string by the function execution sequence. In this way we can filter out redundant functions, and to a certain extent, reduce the running time of RKR-GST algorithm. The experimental result demonstrates that the improved method has achieved the desired goal and the plagiarism detection has a significant effect on the program code. Inadequacies of the replacement tag string have no influence on the same set of rules. If we can ensure a more effective replacement rule, then the final tag string can reflect the system's structural information in a better manner. In conclusion, it is believed that the proposed method in this paper will be useful for theoretical research or practical applications of such systems development for future purposes.

References

1. Berghel, H.L., Sallach, D.L.: Measurements of Program Similarity in Identical Task Environments. ACM SIGPLAN Notices 19(8), 65–76 (1984)
2. Halstead, M.H.: Elements of software science. North Holland, New York (1977)
3. Szuskiewicz, M.: Automatic Plagiarism Detection in Software Code. Information and Communications Tectnology (May 2003)
4. Yamamoto, T., Matsushita, M., Karmiya, T.: Measuring Similarity of Large Software Systems Based on Source Code Correspondence. In: Draft, Division of Software Science, Graduate School of Engineering Science, Osaka University, pp. 4–5 (2002)
5. Ullman, J.D., Aho, A.V., Hirschberg, D.S.: Bounds on the Complexity of the Longest Common Subsequence Problem. In: Proc. of the 15th Annual Symposium on Switching and Automata Theory (swat 1974), vol. 23, pp. 104–109. IEEE Computer Society, Los Alamitos (1976)
6. Gilleland, M.: Levenshtein Distance, in Three Flavors, `http://www.merriampark.com/ld.htm`
7. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 2nd edn. The MIT Press, Redmond (2001)

Computational Prediction of Turbulence Modification in Gas-Solid Two-Phase Flows

Fusheng Yan[1], Jicheng He[2], and Ruquan Liang[2]

[1] Resarch Institute, Northeastern University, P.O. Box318, Shenyang, China 110004
[2] Key Laboratory of National Education Ministry for Electromagnetic Process of Materials, Northeastern University, Shenyang, China 110004
Fusheng_Yan@mail.neu.edu.cn

Abstract. This investigation is concerned with developing a model of the processes involved in turbulence modification in dilute gas-solid turbulent flows. The underlying formulation scheme employs an Eulerian-Lagrangian reference frame, i.e. the carrier phase is considered as a continuum system, while the trajectories of individual particles are calculated using a Lagrangian frame work. A random walk model is used to solve the particle motion equation. The proposed model, along with turbulence modulation models from the literature, is used to simulate a particle-laden vertical pipe flow. The simulation results show that the new model provides improved predictions of the experimental data.

Keywords: Computational fluid dynamics, Turbulence, Simulation, gas-solid flow.

1 Introduction

Computational Fluid Dynamics (CFD) has become a useful means for studying multiphase engineering flows. An important and relevent area where CFD is emerging is in a particle-laden turbulent flow, where the interaction between the carrier phase and the particle phase is complex. The degree of coupling (i.e. one-way or two-way) between the particulates and the gaseous phase is dependent on the amount of particles present [1]. It is well documented that the presence of even small amounts of particles may cause a modification (an increase or a decrease) in the turbulence level or structure of the carrier phase [2]. Experimental data has shown that the ratio of particle diameter to turbulent integral length scale was a crucial parameter for determining turbulence modulation by the presence of particles. The particles suppress turbulence if the ratio is less than 0.1 and the opposite effect is observed if the ratio is greater than about 0.1 [3].

During the last two decades, many researchers have presented "source term models" to predict turbulence modulation [4][5][6]. For these models, an extra term is included in the fluid momentum equation to account for the effect of particles on the fluid phase. The resulting additional source terms to the turbulent kinetic energy and the dissipation rate equations are responsible for turbulence modulation. However, the

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 134–141, 2011.

deficiencies of the conventional source term models lie in two areas: (i) Inability to predict an increase in the turbulence intensity of the carrier phase (ii) Failure to consider the particle crossing trajectory effect. This remains a motivation for the present study.

In this work, the following three mechanisms are considered: (i) Energy transfer through the drag force: this interaction normally results in a decrease in the turbulence intensity of the carrier's phase. (ii) The extra turbulence production due to the wake effects behind particles. (iii) Crossing Trajectory Effect (CTE): It refers to the phenomenon that a particle may fall through the eddy due to a large relative mean velocity between the particle and the turbulent eddy.

2 Governing Equations

A Lagrangian reference frame is employed to calculate the particle trajectories, while an Eulerian reference frame is used to solve the carrier phase equations. Details on the modelling of the two phases are provided below.

2.1 Particle Phase

Particle trajectories are calculated through solution of the particle motion equation. Under the assumptions of dilute flow and very low gas-to-particle density ratios (on the order of 10^{-3}), the particle equation of motion reduces to:

$$\frac{d\vec{u}_p}{dt} = \frac{f}{\tau_p}\left(\vec{u}_g - \vec{u}_p\right) + \vec{g} \tag{1}$$

Where $\vec{u}_g$ and $\vec{u}_p$ are the instantaneous velocities for the gas phase and particle phase, respectively; $\vec{g}$ is the acceleration of gravity; and f is a correction factor to account for the effect of the drag force for flows beyond the Stokes' regime; the symbol τ_p denotes the particle response time. Since the instantaneous gas-phase velocities along the particle trajectories are required to render Eq. (1) solvable, the Stochastic Separated Flow (SSF) model [7] was employed in this work to account for the fluid turbulence through random sampling of the instantaneous gas-phase velocity.

2.2 Gas Phase Equations

Assuming that the mean flow of the gas phase is steady, incompressible, isothermal, and Newtonian with constant properties, then the time-averaged equations for conservation of both mass and momentum have the form in tensor notation below:

$$\frac{\partial U_{gi}}{\partial x_i} = 0 \tag{2}$$

$$\rho_g U_{gj}\frac{\partial U_{gi}}{\partial x_j} = -\frac{\partial \bar{P}}{\partial x_i} + \mu\frac{\partial^2 U_{gi}}{\partial x_j \partial x_j} - \rho_g\frac{\partial \overline{u'_{gi}u'_{gj}}}{\partial x_j} - \bar{F}_{pi} \tag{3}$$

Where the over-bar denotes a time-average quantity; ρ_g is the density of the gas phase; P is the static pressure; U_{gi} and u'_{gi} are the mean and fluctuating gas-phase velocities in the x_i-direction ($i = 1, 2, 3$) respectively; and F_{pi} represents the force exerted by the fluid on the particles per unit volume of mixture. Note that the solid volume fraction is not included since the particulate flow involved in this work is dilute. The two-equation k–ε model is used to solve for the gas-phase turbulence due to its simplicity and robust performance. In addition, the effect of the particles on the turbulent characteristics can be incorporated into the modified k–ε model with relative ease. The particle source terms to the turbulent kinetic energy and the dissipation rate equations can be obtained by starting with the instantaneous gas-phase momentum equation with the extra term included and following the standard procedure for deriving the k and ε equations in a single phase flow. It is noted that the extra particle source terms to the k and ε equations usually appear as a sink, and thus can normally account for the turbulence reduction mechanism due to the drag force between the two phases. To reflect the extra turbulence production mechanism which is attributed to the particle wake effects in this investigation, additional particle source terms must be added to the k and ε equations.

2.3 Modelled Forms for Particle Source Terms

The first particle surce term was modelled based on the model developed by Lightstone and Hodgson where the particle crossing trajectory effect was accounted for [4].

The formulation of the second particle source terms is essentially based on the following two physical observations: (i) the source of the turbulent kinetic energy is supplied via the mean flow field; (ii) the wake formation only occurs for the non-Stokes' regime. It is, thus, assumed that the actual mean force exerted by the particles on the gas phase ($\overline{F_p}$) may be divided into two components such that the first component stands for the particle Stokes' drag force, and the second one represents the difference between the actual mean force exerted by the particles on the gas phase and the particle Stokes' drag force, which contributes to the generation of particle turbulent wakes. It is thus conjectured that the work done by the latter force acting over a distance may account for the energy consumed to produce turbulent wakes behind particles. The resulting second particle source term to the turbulent kinetic energy equation has the form below:

$$S_{kp2} = \left| U_{gi} \frac{\alpha}{\tau_p} (U_{gi} - U_{pi})(f - 1) \right| \tag{4}$$

$$S_{\varepsilon p2} = C_{\varepsilon w} \frac{\varepsilon}{k} \left| U_{gi} \overline{F_{pi,net}} \right| \tag{5}$$

Where $C_{\varepsilon w}$ is a model constant determined through optimizing predictions with experimental data (the subscript "w" implies wake).

3 Test Cases

The experiment of Hosokawa et al [8] is employed for the validation of the simulation results. In their experiment three types of particles (polystyrene, glass, ceramic) with different densities were used. Air was supplied from a blower and moved in the upward direction in a vertical pipe of 5m in length and 30.0mm in diameter. Spherical solid particles were released into the flow at the bottom of the pipe. The measurements were made in the fully developed region of the pipe by using Laser Doppler Velocimeter (LDV). Three test cases were selected in this investigation as detailed in Table 1.

Table 1. Test cases for pipe flow (Hosokawa et al, 1998)

Quantity	Case #1	Case #2	Case #3
Particle Diameter (μm)	900	910	990
Loading Ratio	4.6	4.6	4.6
Particle Density (kg/m^3)	2500	980	3600
Particle Mass Flow Rate (kg/s)	3.9x10^{-2}	3.9x10^{-2}	3.9x10^{-2}
Particle Volum Fraction	2.3x10^{-3}	5.8x10^{-3}	1.6x10^{-3}

4 Numerical Simulation

A finite element based finite volume method was used to solve the gas-phase equations in a Cartesian coordinate system using a colocated grid. Initially, the flow field is calculated without the presence of the particulate phase until a converged solution is obtained. Then a large number of particles are tracked and the source terms to the gas-phase equations are calculated. Thereafter, the gas-phase equations are solved with the particle source terms included. This procedure is repeated until final convergence (a residual of 10^{-4} or less was imposed for this investigation). The simulation was carried out using the commercial CFD code CFX-TASCflow, where a modifiable source code subroutine was used to implement the particle source terms to the k and ε equations.

A wedge-shaped grid was used to simulate this axisymmetric flow as shown in Fig.1. The wedge has an internal angle of 15^o. The radius of the wedge is 15.0mm, and the length of the physical domain is 3m which is long enough to obtain fully developed conditions. The simulation results were taken from the outlet region. A total of 300 nodes are distributed along the axial direction, 20 nodes along the radial direction with more nodes located near the wall, and 3 nodes in the azimuthal direction. To establish grid independence, the number of nodes distributed in the original grid was refined by doubling the number of nodes in the axial and radial directions (600x40x3). The predicted mean axial velocity radial profiles as well as the turbulent kinetic energy radial profiles are virtually identical for the two grids.

The boundary conditions for the gas phase are described as follows: a no-slip condition is assigned to the curved wall. On the axis and the radial planes, perfectly reflecting symmetry boundary conditions are imposed. The inlet velocity was specified

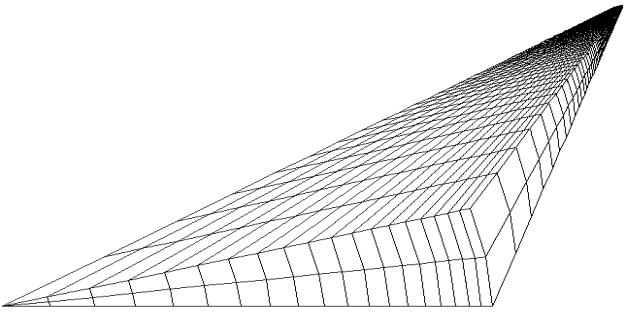


Fig. 1. Computational grid (300x20x3)

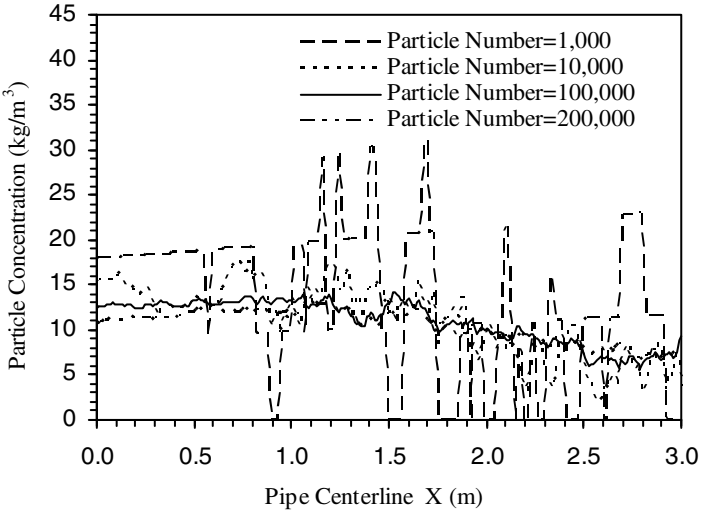


Fig. 2. Predictions of particle concentration along pipe centreline for Case #1

as constant across the inlet plane. The inlet turbulence intensity was specified as 6%; and the eddy length scale was set to 5% of the inlet characteristic length scale. The reference pressure was set to 0 at a single control volume face on the outlet plane. In addition, scalable wall functions were used to resolve the large gradient in velocity near the wall region. The particles were injected with a constant mass flux across the inlet plane. The particle initial velocities were specified to be the same as the gas velocity at the inlet. A uniform particle concentration, consistent with the experiments, was applied. The particles are assumed to perfectly reflect at the walls with a coefficient of restitution equal to 1. The total number of particles injected was 100,000, which was tested to ensure statistical significance by comparing to the results with 200,000 particles as shown in Fig.2.

5 Results and Discussion

The first case, Case #1, considers the polystyrene particles of 900 μm in diameter with a density equal to 2500 kg/m^3. The loading ratio is 4.6. With the mean velocity profile, the experimental data suggests that the radial velocity distribution of the gas phase should become flatter and the velocity gradient in the vicinity of the pipe wall become larger when the particles are introduced into the flow (Fig.3). These tendencies were predicted well by the new model, while the models by CW, MM, LH (short for Chen and Wood, Mostafa and Mongia, Lightstone and Hodgson, respectively) predict a velocity profile which is indistinguishable from that of the clean flow.

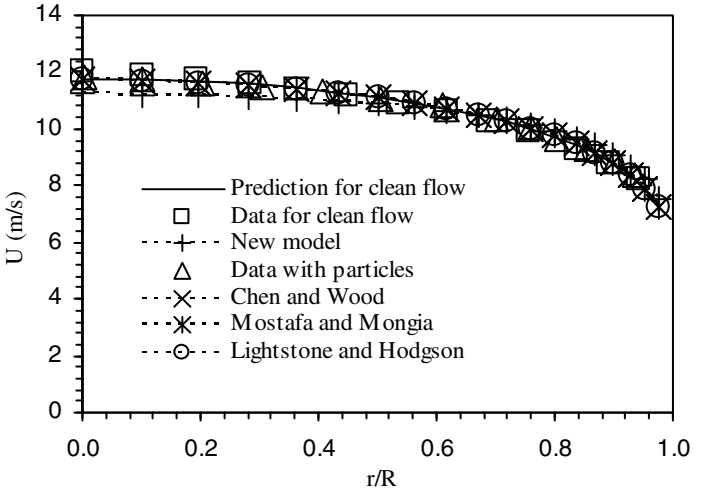


Fig. 3. Predictions of axial mean velocity for Case #1

For the turbulence intensity profile shown in Fig.4, the experimental data shows that the turbulence intensity is increased near the central region of the pipe and decreased away from the inner region with the presence of particles. The prediction given by the new model is seen to be in reasonable agreement with the experimental data in the center of the pipe, but underpredicts near the wall. The other models (CW, LH, MM) significantly under-predict the data since they predict a decrease in the turbulence intensity relative to the clean flow and again give nearly the same predictions. This results from the new model are encouraging because the new model is able to reflect the extra turbulence production with the introduction of large particles.

The similar trends were seen for Case#2 and Case#3. The results were omitted here due to the length of this article.

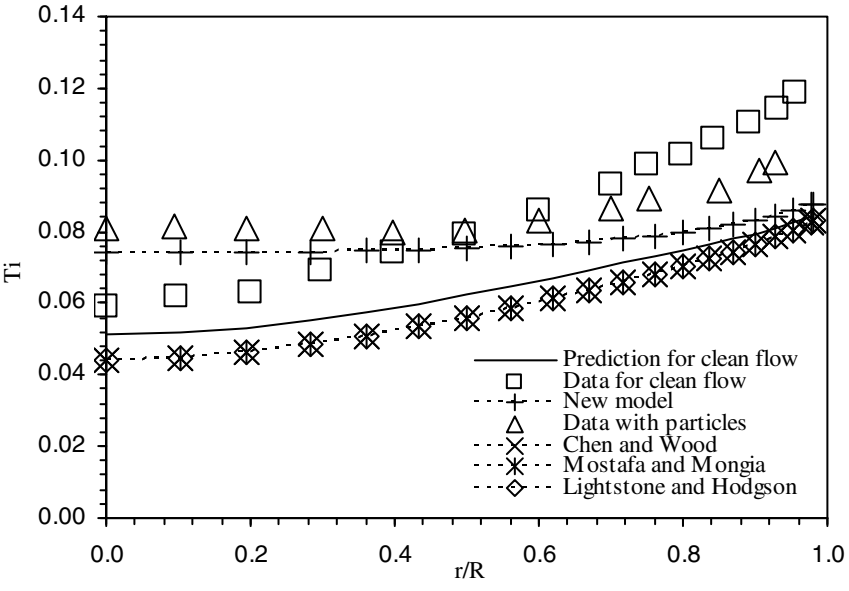


Fig. 4. Predictions of turbulence intensity for Case #1

6 Conclusions

a new mathematical model was presented to account for the turbulence modulation by the presence of relatively large particles in dilute particulate flows. The model together with existing models in the literature was tested against the experimental data of Hosokawa [8]. The results have demonstrated that the new model is able to capture the turbulence enhancement occuring for particle laden flows with relatively large particles. This represents an improvement over predictions obtained using existing models from the literature.

Acknowledgements. The authors wish to thank ANSYS-CFX for providing the CFX-TASCflow computational fluid dynamics code for this research. The financial support from the Northeastern Unversity Research Funding is also acknowledged with gratitude.

References

1. Elghobashi, S.: On Predicting Particle-Laden Turbulent Flows. Applied Scientific Research 52, 309–329 (1994)
2. Chen, C.P., Wood, P.E.: A Turbulence Closure Model for Dilute Gas Particle Flows. Can. J. Chem. Engr. 63, 349–360 (1985)
3. Crowe, C.T.: On Models for Turbulence in Fluid-Particle Flows. Int. J. Multiphase Flow 26, 719–727 (2000)

4. Lightstone, M.F., Hodgson, S.M.: Turbulence Modulation in Gas-Particle Flows:Comparison of Selected Models. Canadian J. of Chem. Eng. 82, 209–219 (2004)
5. Mostafa, A.A., Mongia, H.C.: On the Interaction of Particles and Turbulent Fluid Flow. Int. J. Heat Mass Transfer 31(10), 2063–2075 (1988)
6. Yan, F., Lightstone, M.F., Wood, P.E.: Numerical Study on Turbulence Modulation in Gas-Particle Flows. Heat and Mass Transfer 43, 243–253 (2007)
7. Shuen, J.S., Chen, L.D., Faeth, G.M.: Evaluation of a Stochastic Model of Particle Dispersion in a Turbulent Round Jet. AIChEJ 29, 167–170 (1983)
8. Hosokawa, S., Tomiyama, A., Morimura, M., Fujiwara, S., Sakaguchi, T.: Influence of RelativeVelocity on Turbulence Intensity in Gas-Solid Two-Phase Flow in a Vertical Pipe. In: Third International Conference on Multiphase Flow, Lyon, France (1998)

On Teaching Secretary's Positive Effect in Graduation Thesis Administration

Yan Ma[1], Liping Zheng[1], Shuanglin Jiang[1], and Yiming Fang[2,*]

[1] Tianmu School, Zhejiang A & F University, Lin'an City,
Zhejiang Province, 311300, China
{mayan_ilone,zhenglipingb1517,bjiangsl55}@126.com
[2] School of Information Engineering, Zhejiang A & F University, Lin'an City,
Zhejiang Province, 311300, China
zjfcfym@gmail.com

Abstract. The teaching secretary plays an important role in the administration of graduation thesis as a bridge for communication among schools, guidance teachers and students. Making full use of the teaching secretary's effect is the key to improving the administration of graduation thesis. The paper discusses the teaching secretary's major work in graduation thesis administration and several factors which influence teaching secretary's work are analyzed. Then recommendations are offered for making full use of teaching secretary's effect in graduation thesis administration.

Keywords: Teaching Secretary, Graduation Thesis Administration, Effect, Full Utilization.

1 Introduction

The administration of graduation thesis is always an important part of the undergraduate education management and is characterized by the strong schematization, many link, long duration, multifarious and disorderly content and big workload. The teaching secretary is the executor of the graduation thesis administration and is one of the keys to strengthen the graduation thesis administration. So that, it is of importance to sum up the teaching secretary's major work, analyze the influencing factors and propose the countermeasures.

2 The Teaching Secretary's Major Work

The teaching secretary's major work in the graduation thesis administration, which is basically the same in various universities, mainly consists of intensifying guidance to better topic selection, strengthening the supervisor of graduation thesis writing and standardizing oral defense procedure.

* Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 142–146, 2011.

2.1 Arranging Topics Selection Appropriately

The topics selection must be completed within a short period of time (about 3 ~ 4 weeks). Topics preparation, selection designing, writing plan and opening report, etc, are required to be fulfilled in perfect order within the fixed time span. The detail work mainly includes: making detailed and exhaustive work plan, nominating guidance teacher, holding mobilization meeting, collecting and announcing topics, collecting students' initial intention, paring the guidance teacher and student based on the comprehensive consideration, collecting the selection results and approving the application for writing graduation thesis in other universities or enterprisers.

2.2 Assisting the Supervisor of Graduation Thesis

Through years of practice, it can be observed that many students are busy in seeking employment, graduation practice, internship, taking part in the entrance exams for postgraduate schools or taking an examination of officeholder after the work of graduation topics selection was completed. The graduation thesis writing work is laid aside. Therefore it is necessary to establish a full set of reasonable supervision system. The aims, methods and steps of the graduation research should be examined upon the topics selection. During the thesis writing process, mid-term examination should be taken in good time to see if the progress matches the plan. The attention should be paid to the research results and the amount of work performed before the thesis oral defense. The teaching secretary should give assistance to run the supervision system efficiently. The major work at this stage includes: delivering the notice in time, collecting the students' research results in various stage such as opening report, literature review and translation, collecting the forms filled by the guidance teacher such as mid-term check list and thesis score sheet, reporting emergencies to the department supervisor without delay, summarizing and publishing the progress.

2.3 Strengthening Information Exchange and Intensifying Coordination

Because graduates and many facts affecting the graduation thesis work change in every year, the guidance teacher and the student may not be familiar with the relevant policies, documents, regulations and operational procedures. Therefore the teaching secretary should inform the guidance teacher and the student of the current policies, the plan about the graduation thesis work and the instructions from the schools and superior authorities. At the same time, the teaching secretary should summary the progress and report to the dean accurately.

2.4 Participating Constructively in Thesis Oral Defense

During the thesis oral defense, the teaching secretary ought to take multiple measures to raise the efficiency. Firstly, the students should be reasonable divided into several groups based on the topics. The group should be made up of the students of the same research fields and each dissertation committee is met by the teachers with the same or similar research interests. Secondly, the teaching secretary must arrange a convenient time for the thesis oral defense. Maybe some students have settled into their new position and some of the others are attending an interview for an

employment or master degree candidates out of town. Therefore it may be more reasonable to set the thesis oral defense at the weekend. Thirdly, the teaching secretary need post the notices about time, place and procedure of the thesis oral defense and remind all the students. Finally, the teaching secretary need hear the oral defense randomly and pay attention to whether the time is reasonable and the students' preparation is sufficient in order to give support to the work in future.

3 Influential Factors in Teaching Secretary's Work

3.1 Graduation Thesis's Inherent Characteristics

The implementation time and place differ greatly with the class teaching in the following aspects: Firstly, each guidance teacher is to decide the time and place, which are very flexible, to meet each other in the light of their actual conditions. It's difficult to obtain satisfactory supervisor using conventional management methods. Secondly, with the increasing of employment pressure, more and more students would like to write the graduation thesis in the employing unit with the aim that they can make preparations for their future job in advance as well as know the employing unit deeply. However that means difficulties for the graduation thesis administration due to the employing unit's long way from the school, spreading area and diverse situations. Worse still, if conditions were changed, the administration would be significantly more difficult.

3.2 Lack of Appreciation

Comparing to other posts in the graduation thesis teaching, the work of the teaching secretary takes more time and requires more effort. Even so, the hard work is often misunderstood. There is a widespread belief that the teaching secretaries just perform some simple task such as: posting up notice, collecting data, recording score, etc. It is so easy that everyone is capable of it.

4 The Creation of Favorable External Conditions and Environment

The management level, working condition and management effect of the teaching secretary directly influence the progress of graduation thesis teaching in order as well as the quality of the graduation thesis. The leaders at all levels and relevant department in college and university should realize the importance of the teaching secretary in the graduation thesis administration and create favorable external conditions and environment. That is of great significance to the graduation thesis teaching.

4.1 Setting up a Scientific Innovation and Assessing System

Considering the lasting time span and burdensome task of the teaching secretary's work in the graduation thesis administration, a scientific innovation and assessing system should be set up to ensure the benefits of those working silently and hard. The evaluation results should be based on the following aspects: attitude to work, work

skill, work load and achievement. The public citations and material rewards should be favored for teaching secretary that has a strong sense of responsibility and high working efficiency. Whereas those, lacking administrative ability and often trying to evade the question, should be ordered to make corrections within a time limit, or taken disciplinary measures. The criteria for rewards and penalties must be perfectly clear so as to make full use of the human resources and mobilize the enthusiasm of the teaching secretary to improve their service quality.

4.2 Upgrade the Pay and Status of Teaching Secretary

For the construction of the professionalization on the teaching secretary in colleges and universities, an important question that must be settled is how to ensure the stability of the secretary staffs and how to attract the candidate with the characteristics of education background, confidence, creativeness and go-getter to join the staffs. This can be viewed from two aspects. Firstly, restructure reward system to enhance the effect of teaching secretaries. At present, the salary and post allowance in colleges and universities are linked with the professional title or the positions. This leads to that the income of the teaching secretary is significantly lower than other staff in the colleges and universities. So the peculiarities of the specific group should be considered. When resources permit, an extra allowance can be added. Secondly, the school should take special consideration and provide some assistance in promotion, determining technical or professional titles or in other aspects.

5 Enhancing the Teaching Secretary's Own Quality

5.1 Strengthening the Service Consciousness

As the organizer and executor at the grass-roots level, the secretaries' position determines their main duty is to serve the dean, guidance teachers and the students. With the constant deepening of the reform of university management system, the teaching secretary should strengthen the service consciousness and adopt the idea of "thinking about the interests of the teachers and students, serving for teaching". The core of the work is teachers and students. The position and role of teachers and students in the teaching link of graduation thesis must be recognized profoundly. The secretary should serve the graduation thesis work earnestly and sincerely so as to improve the quality of service and create a favorable environment for the graduation thesis teaching.

5.2 Improving Mental Qualities

The psychological imbalance is very common in the colleges and universities partly due to the income gap. Then some people would do just enough work to get by, believing "so long as one remains a monk, one goes on tolling the bell". This is an abnormal phenomenon which must be eradicated. The graduation thesis administration is a high professional work, and it requires the teaching secretary have good administrative skill and steady mentality, especially dedication. Smile, sincerity and patience are essential qualities. The teaching secretaries should think about the

guidance teachers and the students at all times and live in harmony with others. Misunderstanding and grievances often come up in work. The secretaries must learn to be tolerant and always suffer imbalance gladly. Outstanding contribution stems from working perseveringly, sparing no efforts and being confident when facing any difficulty.

5.3 Mastering Modern Office Management Skill

The graduation thesis administration work is very technical and practical with the characters of big workload and weight responsibility. It requires teaching secretaries have great enterprise and consciousness along with modern office management skills. The teaching secretaries should be skillful in MS Word, Excel, PPT, Outlook and other software and promote the communication by new channel such as QQ group, E-mail group sending, blog, microblogging, etc. PC skills are the important components of graduation thesis administration modernization as well as the inherent requirement of the improvement of the graduation thesis quality.

6 Conclusions

As the organizer and executor at the grass-roots level, the teaching secretaries are always in the forefront. They participate in management, perform the tangible work, work as bridges among the superior authorities, department, the guidance teachers and graduates and continuous develop their serving function in order to improve the quality of graduation thesis. The role of the teaching secretary in graduation thesis administration is indispensable. The leaders at all levels and relevant department in university and college should strengthen the construction of teaching secretary. Effective measures should be taken according to actual conditions with the aim to make full use of the secretary's positive effect in graduation thesis administration.

Acknowledgments. This work acknowledges the support from the Key Project for Education Reform Research of Tianmu School of Zhejiang A & F University (TMDZ0802).

References

1. Xie, X.M.: The Role of Teaching Secretary in Graduation Thesis Administration. Chin. J. Beijing Educ. 9, 43–44 (2009)
2. Liu, L.: Analysis on the Role of Teaching Secretary in Improving Teaching Quality–6Years Teaching Management Experience for Adult Education. Chin. J. Data Cul. Educ. 20, 107–108 (2006)
3. Xu, Y.H.: Enhancing Procedure Administration of Undergraduate Thesis and Project. Chin. J. Zhejiang Wanli Univ. 20, 151–153 (2007)
4. Ren, L.Q.: On Specialization of Teaching Secretary in College and University. Chin. J. Ningbo Univ. 30, 137–139 (2008)
5. Xiao, W.: Reflections on a University Teaching Secretary's Work in the New Era. Chin. J. Educ. Mod. 85, 75–80 (2008)

Back Propagation Neural Network on the Forecasting System of Sea Food Material Demand

Chih-Yao Lo

Department of Leisure Management, Yu Da University, No 168, Hsueh-fu Rd, Tanwen Village, Chaochiao Township, Miaoli County, 36143 Taiwan (ROC)
jacklo@ydu.edu.tw

Abstract. The price of fish product is changing daily based on the market demand and the supply from feeding and fishing at any moment. We find that the production and demand of fish product are different in different seasons based on the investigation and analysis on the fish product market. Therefore it will be much advantageous for the operating personnel if the short-term demand of customer in the future can be forecasted. This research makes use of the back propagation neural network algorithm to forecast the fish product demand, so that the order demand in the future can be forecasted on the base of the existing order data. The object is to improve the competitive force of industry and maximize the profit.

Keywords: Fishery, Back Propagation Neural Network, Forecasting.

1 Introduction

The transportation and sale of fish product means the economic activity that the fish product is sent to the consumer or subsequent manufacturer through the working procedure from production to collection, classification, transportation, storage, processing, sale, financing and market information collection. However, the price is an important factor with great influence on the operating other than the product freshness in the fish product transportation and sale industry. The price of fish product will reflect in the fish product auction market based on the market demand and the supply from feeding and fishing at any moment. Therefore it will be much advantageous for the operating personnel if the short-term demand of customer in the future can be forecasted [1].

This research achieves the object to forecast the fish product demand with back propagation neural network algorithm. We also expect to combine the theory and practice with the cooperation with the fishing personnel in the future [2-3].

2 Literature Review

The back propagation neural network is one of the most representational modes used extensively in all neural network learning modes at current. Werbos and Parker have brought forward the basic concept in 1974 and 1982 respectively, but it is till 1985

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 147–154, 2011.

that Rumelhart, Hinton and Williams of Stanford University brought forward the propagation learning rule or generalized delta learning rule that this theory and algorithm has been defined definitely [4][5].

The basic principle of back propagation neural network is to make use of the gradient descent method to minimize the error function and thus derive the delta rule, and its idea is to reduce the difference between the actual output and expected output with a successive correctional value. From the viewpoint of mathematics, the correction of synapse value is direct proportion to the 1st differential of error value, which can testify it will converge to a stable state in the process of network learning, equivalent to the minimum value of a curve on the plane[6].

The processing process of back propagation network includes forward pass and backward pass, and the error can be reduced and the expected learning fruit can be achieved through these two stages. The forward pass begins from the input layer, transfers the original data to the vector and input into the network, and they will be calculated by each layer of nerve cell till the last layer of network. The backward pass is to pass from the output layer to the input layer and consisted of the error calculation and weight value update. Its method is to compare the difference between the object output and actual calculated values at first, then adjust the synapse value of network to reduce the error to the minimum[7-8].

There are 2 works for each nerve cell:

1) Connection

$$u_i(k+1) = \sum_{j=1}^{N} W_{ij}(k) a_j(k) + \theta(k) \tag{1}$$

2) Activation

$$a_i(k+1) = f(u_i(k+1), u_i(k), a_i(k)) \tag{2}$$

The nerve cell will implement the connection and activation repeatedly to update the activation value on the nerve cell. It will be passed to the other nerve cells when a new activation value emerges. Equation (1) and equation (2) is named system dynamical equations together, in which index k indicates the update times, represents the output value of nerve cell i , w_{ij} denotes the weight value between nerve cell i and j , θ_i is the internal threshold value of nerve cell i , N is the number of nerve cells connected with nerve cell i , $f(\bullet)$ is the transition function, which is from the simplification of biologic effect in the nerve cells.

3 Methodology and Steps

According to the relevant literatures, the supervised learning network must get the training sample in the field of problem, including the input value and output value. This network will be trained at first, then the weight value will be adjusted according to the difference between the actual output value and expected output value; it means that the network will learn the map rule between these two values and then apply it to

the new case. The object of supervised learning is to reduce the difference between the actual output value and expected output value, and it can be used in forecasting and classification. Therefore the famous back propagation neural network algorithm in the supervised learning methods can be used to forecast the fish product demand in the future, and the application method will be detailed as follows.

3.1 Demand Forecasting Parameter Conversion

We find that the production and demand of fish product are different in different seasons based on the investigation and analysis on the fish product market. Firstly, we convert the fish product order of each cycle (solar term, week or month) into the input matrix ($X_1 \ldots, X_n$) according to the amount and gross amount of different kinds of fish product, and the output matrix will be obtained from the proper conversion of the amount of different kinds of fish product in the order of next circle. For the conversion method, refer to Table 1:

Table 1. Comparison of Input Matrix Conversion

Input Parameter	Meaning
X_1	Gross Weight of Demand (Kg)
X_2	Demand Ratio of Fish Product A
X_3	Demand Ratio of Fish Product B
X_4	Demand Ratio of Fish Product C
X_n	Demand Ratio of Fish Product N

3.2 Demand Forecasting Algorithm

The back propagation neural network order demand forecasting module will be built according to the back propagation neural network algorithm after the proper input matrix and output matrix is obtained, and the module structure will adopt a hide layer, see Fig. 1, which can also be denoted with equation (1):

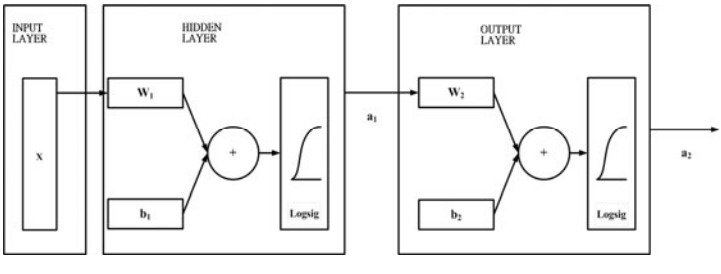


Fig. 1. Back Propagation Neural Network Order Demand Forecasting Structure

In which X in the input layer indicates the input matrix, and this matrix is the parameter matrix conversed from the order of each cycle according to the amount of different kind of fish product. a_1 is the output matrix of hide layer, a_2 is the output matrix of output layer, which represent the order demand matrix to be forecasted.

$$a_2 = f(W_2 \bullet f(W_1 \bullet X - b_1) - b_2) \tag{3}$$

In which $f()$ indicates the conversion function, the W_1, b_1, W_2, b_2 in the hide layer and output layer will be obtained from the calculation of back propagation neural network training algorithm, which will be detailed below. According to the relevant literature, the conversion function expected to be used in this research is Logsig, and its output and input relationship is as showed in Fig. 2:

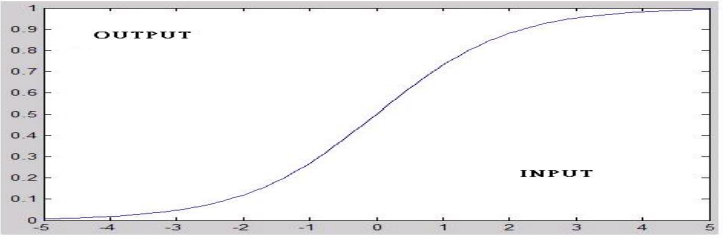


Fig. 2. Conversion Function Input and Output Relationship

The learning and training process of weight matrix of neural network hide layer and threshold vector , weight matrix of neural network output layer and threshold vector is detailed as follows:

1) Input a set of training sample vector X and expectation object vector T.

2) Calculate the actual output vector a_1 of network.

1. Calculate the output vector a_1 of hide layer.

$$n = \sum W_1 \cdot X - b_1 \tag{4}$$

2. $$a_1 = f(n) = \log sig(n) = \frac{1}{1+e^{-n}} \tag{5}$$

3. Calculate the actual output vector a_2.

$$m = \sum W_2 \cdot a_1 - b_2 \tag{6}$$

$$a_2 = f(m) = \log sig(m) = \frac{1}{1+e^{-m}} \tag{7}$$

3) Calculate the difference E between the output vector a_2 and object vector T.

$$E = T - a_2 \tag{8}$$

4) Update the weight value W and threshold value b.

1. Update the weight value W_2 and threshold value b_2 of output layer.

$$W_2(new) = W_2 + \Delta W_2 \tag{9}$$

2. $$b_2(new) = b_2 + \Delta b_2 \tag{10}$$

3. Update the weight value W_1 and threshold value b_1 of hide layer.

$$W_1(new) = W_1 + \Delta W_1 \tag{11}$$

4. $$b_1(new) = b_1 + \Delta b_1 \; b_1(new) = b_1 + \Delta b_1 \tag{12}$$

5) In which the calculation method of weight correction value ΔW and threshold correction value Δb is as follows:

$$\Delta W_2 = \eta \delta_2 a_1^T \tag{13}$$

$$\Delta b_2 = -\eta \delta_2 \tag{14}$$

$$\Delta W_1 = \eta \delta_1 X^T \tag{15}$$

$$\Delta b_1 = -\eta \delta_1 \tag{16}$$

η in the equation (13) to (16) represents the learning efficiency, which is set by the user, and the calculation method of δ_1, δ_2 is as follows:

$$\delta_2 = -2F_2(m)(T - a_2 \tag{17}$$

$$F_2(m) = \begin{bmatrix} \cdot f_2(m_1) & 0 & 0 \\ 0 & \cdot f_2(m_2) & 0 \\ \cdot & \cdot & \cdot \\ 0 & 0 & \cdot f_2(m_i) \end{bmatrix} \tag{18}$$

$$\cdot f_1(m_i) = \frac{d}{dn}\left[\frac{1}{1+e^{-m_i}}\right] = (1 - a_2)(a_2) \tag{19}$$

$$\delta_1 = F_1(n)(W_1)^T \delta_2 \tag{20}$$

$$F_1(n) = \begin{bmatrix} \cdot f_1(n_1) & 0 & 0 \\ 0 & \cdot f_1(n_2) & 0 \\ . & . & . \\ 0 & 0 & \cdot f_1(n_i) \end{bmatrix} \tag{21}$$

$$\cdot f_1(n_i) = \frac{d}{dn}\left[\frac{1}{1+e^{-n_i}}\right] = (1-a_1)(a_1) \tag{22}$$

Repeat the above steps till the error E has not obvious changed, i.e. the convergence is achieved and the learning is finished. The deposited weight value W_1, W_2 and threshold value b_1, b_2 will represent the order feature concluded from the former orders after the learning is finished. Substitute the weight value matrix and the threshold value matrix into the equation (1) for the subsequent back propagation neural network order demand forecasting module to calculate the order demand forecasting.

4 Validation and Results

The simulation of this paper is based on a quarter order amount in 2009 for the traditional fishery product supplier A.

The test data that is provided by the supplier A can be classified into the training samples and the test samples, and the corresponding sample sizes are 30, 30, 45, 30, 60, and 30 respectively. The simulations are performed based on the algorithm of Back Propagation Neural Network, and the simulation results are summarized as Table 2.

Table 2. Estimation Result

Training					Testing		
Sample size	Success training	Success rate	Mean square error	Learning times	Sample size	Success Testing	Success rate
30	60	100%	10^{-15}~10^{-16}	37481	30	15	50%
45	72	100%	10^{-15}~10^{-16}	4089	30	19	63%
60	90	100%	10^{-3}~10^{-4}	100000	30	24	80%

Convergence process of Mean square error in these simulations are showed as Fig. 3-5.

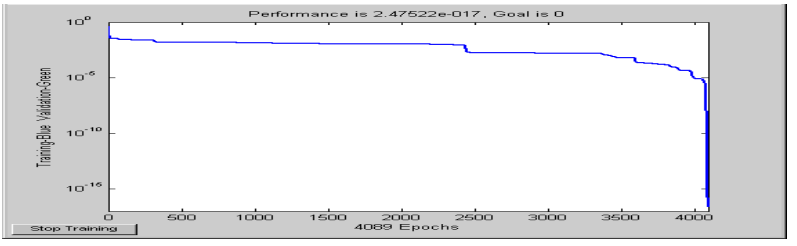


Fig. 3. Convergence process of sample size 30

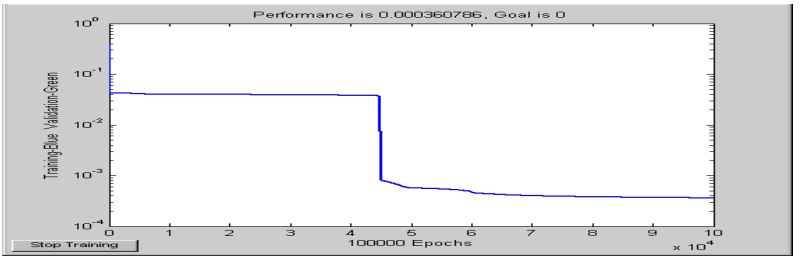


Fig. 4. Convergence process of sample size 45

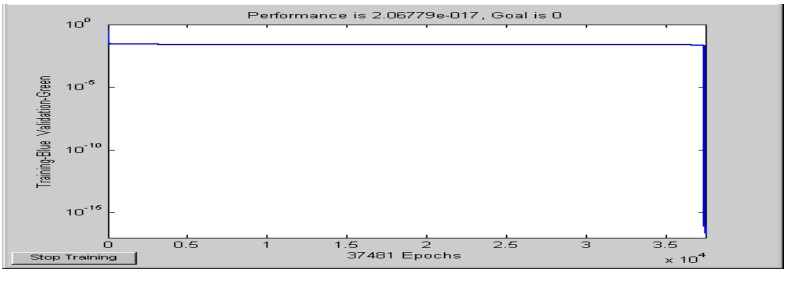


Fig. 5. Convergence process of sample size 60

5 Conclusions

According to simulation results, shown as table 2, it is applicable for Back Propagation Neural Network on the estimated system of fishery product demand. However, because it is difficult to collect historical data from traditional fishery industry, more test data will be discovered for performing training and improving the success rate of estimation that can be used to modify the proposed algorithm for achieving higher efficiency in the future.

This research makes use of the back propagation neural network algorithm to attempt to forecast the fish product demand, and the object is to improve the competitive force of industry and maximize the profit. In the future research, we will seek for the cooperation with the leading enterprises in the fish product transportation and sale industry actively, and apply the forecasting method to the former orders of

the enterprises to conclude the feature of order data on the network, then validate the short-term demand forecasting of enterprise in the future via the order feature to achieve the object of combining the theory with the practice.

Acknowledgments. This research is supported by NSC-95-2622-E-412-001-CC3.

References

1. McGarvey, R.: Demand-side fishery management: integrating two forms of input control. Marine Policy 27(3), 207–218 (2003)
2. De, J.G., Antoni, V.A.: Forecasting threshold cointegrated systems. International Journal of Forecasting 20(2), 237–253 (2004)
3. Smith, C.D., Mentzer, J.T.: Forecasting task-technology fit: The influence of individuals, systems and procedures on forecast performance. International Journal of Forecasting (2009) (in press, corrected proof, available online)
4. Wang, J.D., Xiao, D.Q.: Introduction of Neural Network and Fuzzy Control Theory. Quanhua Science & Technology Books Co. Ltd. (2002)
5. Huang, R., Xi, L., Li, X., Liu, C.R., Qiu, H., Lee, J.: Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. Mechanical Systems and Signal Processing 21(1), 193–207 (2007)
6. Min, Y., Wang, Y.J., Cheng, Y.P.: An incorporate genetic algorithm based back propagation neural network model for coal and gas outburst intensity prediction. Procedia Earth and Planetary Science 1(1), 1285–1292 (2009)
7. Wang, T.Y., Huang, C.Y.: Optimizing back-propagation networks via a calibrated heuristic algorithm with an orthogonal array. Expert Systems with Applications 34(3), 1630–1641 (2008)
8. Huang, C.Y., Chen, L.H., Chen, Y.L., Chang, F.M.: Evaluating the process of a genetic algorithm to improve the back-propagation network: A Monte Carlo study. Expert Systems with Applications 36(2), 1459–1465 (2009)

Research on Cache Management for Mobile Learning Devices

Ye Liang

Department of Computer Science, Beijing Foreign Studies University,
Beijing, China
liang_ye@sohu.com

Abstract. Expand the bandwidth of mobile communication devices and enhanced cache utilization are two aspects to improve quality of service for mobile learning, the pace of expansion of mobile communication bandwidth has not stopped, but the pace of development of mobile learning applications and bandwidth will always be consistent with the same expansion rate. The new application of large amounts of data of mobile learning will soon run out of resources in bandwidth reservation with the new technology. In this paper, we put forword a new mobile cache management mechanism, and can improve the the efficiency of information collection, queries, synchronization, caching, prefetching, maintenance for mobile learning devices. Experimental results show that the performance of cache management can service better within a limited area for mobile learning devices.

Keywords: mobile learning, cache management, mobile devices.

1 Introduction

Mobility has become an outstanding characteristic in learning a foreign language field in the world. In recent years, with the improvement of mobile communication rate and the discount of fee for mobile communication, and the enhance of performance of handheld mobile computing, and the increase of storage of mobile equipment, the cell phone, PDA, handheld computers and other handheld mobile devices of the students in their pockets or bag will make it possible to allow the them to process, access, or transmit the knowledge and living information at any time and everywhere. At the same time it is also possible for teaching workers to carry out educational activities with hand-held mobile devices and wireless networks, and it is possible to achieve life-long learning. Therefore, it has become this century's cutting-edge research how to use handheld mobile devices to get better education information, teaching interaction and education at home and abroad of the hot spots.

Mobile learning is a new concept in the background. It refers to learners in his own need to learn any time, any place, communicate to others, and study with others through hand-held mobile equipment and wireless Internet for learning resources. In this background, it is particularly and urgent to research on the mobility of study of foreign language learning. It can be summarized with mobile learning with the similarities and differences of traditional learning, and also can move deeper into the

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 155–162, 2011.

characteristics how to link learning in classroom and learning out classroom. This is an important topic of foreign language teaching filed in China, and it is also an important topic of foreign language teaching filed in collages. [1], [2], [3].

2 Related Work of Mobile Learning

Compared with other parts of the world, it is earlier for North America to develop the mobile learning, and it is studied originally and deeply. Early in this century, American researchers have predicted, the rapid development of hand-held devices and wireless communication technology in the early 20th century will bring out huge influence in the human society. Under this conception, the development of the mobile learning and the mobile teaching with handheld devices is more popular in North America. Because of the social investment and national attention, mobile learning of basic education field is researched and applied deeper in North America. As is known to all, mobile study not only requires considerable funds and resources support, but also need to rethink the integration of teaching and technology. Mobile learning environment makes all students could use all sorts of digital equipment and services to study in anytime and any place. It not only focuses people use a variety of different tools, but also urges study in the learning process of teachers and students actively, can choose suitable for specific tasks need equipment and tools, truly to analyze the information; Through all kinds of ways to create new knowledge, communication to learn on your own income. In north American and Europe, a growing number of schools, area, even nationwide began to support the research and application of mobile learning. In domestic teaching world, the network teaching, self-learning, electronic teaching research more, but put all sorts of mobility learning tools and learning means together in research are rare, and this research on foreign language teaching world is very important and timely.

Mobile learning in the application of the difficulties of teaching mainly reflects in the following aspects:

Because move the device itself is learning, operating inconvenience and small screen wireless network speed slower characteristic, therefore, how to organize the foreign language learning resources, make its intellectual content brief, interface convenient operation, is mobile learning in foreign language teaching application on the most urgent need to address the problem.

People in the "mobile", the height is and attention of scattered "associated", which makes learners in certain "fragmentary" time for learning, relative to the stable environment in a relatively short time study environment, which makes every time the content of study is not a lot. Mobile learning this "pieces" type learning experience in learning resources construction, demanding respect should notice development suitable for spare time of study materials. Therefore, how to foreign language learning resources for a series of refinement decomposition of knowledge, so that learners son to in a short time, the master swiftly and effectively in teaching field is mobile learning another dilemma in application realistic problem.

Mobile learning segment sex also determines the mobile foreign language teaching content must have very good continuity, can around coherence. Mobile learning teaching activities by learners free time is arranged, easy to cause and learners in learning content not integrity and long-lasting. Therefore, how to guarantee knowledge resources can constantly meet learners need and learners consistent level of knowledge, but also mobile learning in foreign language teaching application on issues that need to be resolved.

Mobile learning is not easy to perform various quantitative statistics, measurement and observation, how to establish the corresponding quantitative method, and USES the method of investigation and research on foreign language learning method in the student mobility of overall study life roles is this topic needs to solve one of the key.

3 Index Maintenance and Updating Algorithm

Cache Management is the most important task for Mobile Learning Devices. We set up a CG TPR-tree index to manage the cache. The CG TPR-tree index preserves the basic idea of TPR-tree index design. [4], [5] It builds index for all the moving objects on the basis of R-tree and adds the relative information of corresponding sub-tree to the middle node and the information includes the integral value of the rate's maximum value and minimum value within a group from the reference moment and the pointer to the parent node and so on, thus forming the extended TPR-tree. Its record form is five-group < MBR, $\overrightarrow{\text{Vector}}_{\max}$, $\overrightarrow{\text{Vector}}_{\min}$, ptr_{parent} , ptr_{child} >, separately representing the node's boundary rectangle, the rate's maximum integral value and minimum integral value from the reference time, the pointer point to the parent node and the child node. As the time goes by, every node groups automatically through moving and generates all kinds of attribute information in five-group, preparing for future index maintenance and updating. Initially, the index structure built through the way of R-tree causes the extension of MBR in every node due to bearing the moving objects with all kinds of motion characteristics. To ensure the index efficiency, we put forward GMOM and GMOU o the basis of GG TPR-tree index method.

3.1 Description of GMOM Algorithm

GMOM algorithm maintains GG TPR-tree nodes at all levels by grouping and its idea is as follows: all the grouping moving objects are not possible to move with the same rate and direction which will inevitably lead to the extension of MBR as time goes by and the index efficiency will be lowered. When TPR-tree's MBR is on a certain layer overlap, the GMOM algorithm starts to eliminate the overlapping field. According to the analysis of the moving kinds of the moving objects above, the reasons that causes the extension of MBR is as follows:

Suppose a group of moving objects has the minimum boundary rectangle S_0 in the reference time t=0. At this moment, they do the linear running with the speed v_0

and their velocity projections are v_x and v_y in the two quadrants when they do the orthogonal decomposition. The upper and lower bounds of the boundary rectangle are $v_{x\max}$, $v_{x\min}$ and $v_{y\max}$, $v_{y\min}$ respectively, and equals v_x and v_y.

The rate will inevitably fluctuate slightly while the moving objects are moving. When it happens, the upper and lower bounds of MBR will change, that is to say, $v_{x\min} < v_x$ or $v_x < v_{x\max}$ and then $S_0 < S_T$, thus leading to the extension of the boundary rectangle. The moving objects within the group keep constant motion based on the average speed, so the expansion of the area $\Delta S = S_T - S_0$ is ineffective to the index. When t=T, implement GMOM algorithm and cut this area.

Through inquiring the relative nodes of GG TPR-tree, we can know the parameter number. Then GMOM algorithm can be effectively implemented.

Suppose the speed of a group of moving objects is v_1 in the reference time t=0 and their velocity projections are v_x and v_y in the two quadrants when they do the orthogonal decomposition. The upper and lower bounds of the boundary rectangle are $v_{x\max}, v_{x\min}$ and $v_{y\max}, v_{y\min}$ respectively, meeting $v_x = v_{x\max} = v_{x\min}$, $v_y = v_{y\max} = v_{y\min}$ and they has the minimum boundary rectangle S_0. At this moment, the moving objects within this group begin to reduce the rate in turn from v_1 to 0 with the acceleration a_1. Due to the hysteresis of the stop act, all the moving objects stop when t= T_1, when t= T_2, this group's moving objects speed up to the rate v_2 with the acceleration a_2, when t= T_3, all the moving objects achieve the rate v_2 and keep the constant linear motion with this rate.

In the process above, during the period that $t \in [0, T_1]$, because the moving objects which are at the front of the queue slow down, the moving speed of the lower bounds of the boundary rectangle starts to decline. Due to the hysteresis of the stop act, the moving objects at the back part maintain the original speed. Therefore, the upper bounds of the boundary rectangle keep the same speed and the area of the boundary rectangle expands, bearing invalid index space.

Suppose in the reference time t=0, a group of moving objects' minimum boundary rectangle S_0, rate v_0, velocity projection v_x and v_y, upper and lower bounds $v_{x\max}$, $v_{x\min}$, $v_{y\max}$, $v_{y\min}$ are same as (2), and at the moment, the moving objects turns in the same direction one after another. In the process above, because the movement has the property of hysterisis, there is the adverse impact to the index coming from the extension of the boundary rectangle. A turning can be seen as the deceleration in one direction and acceleration in another.

Select the inefficient index space and delete by orthogonal decomposition and synthesis.

Other situations combined by the three conditions above

According to the analysis of the reason that leads to inefficient index space, here is GMOM algorithm that cuts the inefficient space within the boundary rectangle of GG TPR-tree

Algorithm 1: GMOM
Input: GG TPR-tree root
Output: maintained GG TPR-tree MBR

```
BEGIN
1.  If  root.children • NULL Then
2.            For all child ←  root.children Do
3.              If child.children • NULL
                     Then childstack   child
4.        End For
5.  End If
6.  For all node1   childstack Do
7.        For all node2   childstack Do
8.              If node1.MBR ∩  node2.MBR • NULL
                     Then
9.                    For all child   node2.children Do
10.                         If child.children • NULL
                              Then childstack ←  child
11.                   End For
12.             End If
13.       End For
14. End For
15. For all node1, node2   childstack Do
16.       If node1.MBR ∩  node2.MBR = NULL Then
17.             Delete node from childstack
18.       Else
19.             cut out the node.MBR
20.       End If
21. End For
    END
```

In GMOM algorithm, one to five steps achieve the root node's child nodes from GG TPR-tree, not including the child nodes that belong to the leaf nodes themselves; six to fourteen steps are the recursive search for non-leaf nodes of all levels that cause the overlapping of MBR; fifteen to nineteen steps delete all the sub-nodes that have nothing to do with the overlapping of MBR and cut and contract the rest nodes MBR as the analysis above.

3.2 Discription of GMOU Algorithm

GMOM algorithm groups and updates the nodes of GG TPR-tree and its idea is as follows:

When MBRs at a certain layer of GG TPR-tree overlap as time goes by and GMOM algorithm can not eliminate the overlapping field, we should consider using

GMOM algorithm to re-group the relative nodes of GG TPR-tree. According to the analysis of the moving kinds above, GMOM algorithm can not eliminate the nodes' overlapping field and the reasons are as follows:

Because the task and the direction of the moving objects within a group are different at some moment, they separate gradually. That is to say, when the compatibility of the moving objects reaches to the low point from a higher one, due to the deviating moving direction, the boundary rectangle expand at once and interfere the index efficiency. In this case, we need to split the moving objects within this group.

Two groups of moving objects cross with each other guided by the road network due to the need of the task at some moment. They have the similar location and speed. That is to say, when the compatibility of the moving objects reaches to the high point from a lower one, in order to improve the index efficiency, we need to merge the two groups and index.

Two groups of moving objects cross with each other guided by the road network at some moment. Even though they have the similar location and speed, their moving direction is opposite. That is to say, the capability of the moving objects within this group is low and the grouping of the moving objects remains unchanged in order to ensure the index efficiency.

Other situations combined by the three conditions above.

According to the analysis of the reason that the overlapping of MBR can not be eliminated, GG TPR-tree index re-groups the moving objects and the description of the re-grouping determination and updating algorithm is as follows.

Algorithm 2: GMOU
Input: childstack of GMOM
Output: updated GG TPR-tree indexing

```
BEGIN
1.  For all childstacks Do
2.        GG (childstack)
3.        CG (childstack)
4.  End For
END
```

In GMOM algorithm, the second step calls that GG algorithm carries out the grouping operation to the overlapping nodes of MBR; the third step calls that CG algorithm carries out the merging operation within the group to the overlapping nodes of MBR.

4 Experiment Verification

4.1 Contents and Settings of the Experiment

In order to evaluate GG TPR-tree index and the grouping algorithm and clustering algorithm based on GG TPR-tree index as well as the performance of algorithm on the basis of the grouping maintenance and updating, we carry out the following experiment. We adopt the data of roads above grade three in the 18Km *18Km range in the city proper within Beijing's Fourth Ring Road as our spatial location. The

moving objects data generates at random. We experiment for twenty times and the size of the data set begins from 5000 to 100000 with the increments of 5000 every time, representing the simultaneously moving vehicles on the roads above grade three within the spatial location of the experiment. The moving objects are indicated by coordinate points, randomly distributed on all roads. The traffic lights at the intersections set as 100 seconds: the proportion of 100 seconds to allocate the time of red lights and green lights. To make simpler, we suppose if all turns (including right turns) are controlled by the traffic lights, with all the roads are two-way and three moving objects are able to go side by side in each direction on grade one roads, two are able on grade two roads, and one is able on grade three roads.

The hardware environment of the experiment: CPU: Intel Pentium 1.8GHz; Memory: 256MB; Hard Disk: 40G, 7200RPM. Software environment: Microsoft Visual C++6.0

The simulator performs once per second. The moving range of each object is 0~6 coordinate points, corresponding to an actual speed of 0~65Km/h, with the moving distance as long as possible without the block of any other moving objects. Before moving, it must be observed that the next coordinate point of the moving direction of the objects has no other moving objects, namely any two moving objects have at least an interval of one coordinate point. When the moving object is at the n fork, it will be decided to turn to each fork according to the probability of $p_1, p_2, \ldots\ldots, p_n$ (the probability of going straight marks as p1, and $p_1 + p_2 + \ldots\ldots + p_n = 100\%$). If the moving objects at the fork are full, the turning direction will be re-decided. If the moving objects at all forks are full, they will not move in the implementation.

4.2 Result and Analysis of the Experiment

To make the contrast convenient, we adopt the mode of calculating the access times of index operation to disk I/O to define the performance of index updating algorithm. To the EBUU algorithm of HV TPR tree index, to the maintenance of the index of our moving objects after their updating and their speeds fall into other speed sets, the access times of disk I/O are $total_1 = (h+3) \times I/O$. To the other conditions, the access times of disk I/O $total_2 \geq 0$, the total access times $total \geq total = (h+3) \times I/O_1$, and the total access times of updating algorithm based on the grouping GG TPR tree are the disk I/O cost when a moving objects group and cluster and the cost of (n+1) times of insertion deletion of nodes, that is $total' = (n+1) \times (h+3) \times I/O_1 + a \times I/O_1$.

5 Conclusion

By grouping the moving objects which are similar in location and movement with making use of the grids to realize the whole maintenance and control in the groups, reducing the intermediate links of index maintenance, hereby increasing the maintaining and updating efficiency of tree index. Meanwhile, we bring forth the GG algorithm and the CG algorithm both based on the grid technology, as well as the CG algorithm and the GMOU algorithm both based on grouping. Backed by the index structure based on GG TPR-tree, times of index insertion and deletion dramatically

reduce by means of the algorithms above when maintaining and updating the index of mass data within the limited range, lessening the access frequency of disk pages. Therefore, with adopting the above-mentioned algorithms, the GG TPR-tree performs better at the indexing aspect than the existing indexing method for moving objects

Acknowledgment

This paper is supported by the Fundamental Research Funds for the Central Universities (No.2010XJ017). Without this help, this work would never have been completed.

References

1. Anandharaj, G., Anitha, R.: A Distributed Cache Management Architecture for Mobile Computing Environments. In: Advance Computing Conference, pp. 642–648 (2009)
2. Kumar, C., Kiranmayi, K., Bhargava, A., et al.: Efficient Bandwidth Utilization in Mobile Environments for Cache Consistency Using Queried Data (CCUQD) Approach. In: INDICON, pp. 1–7
3. Sagayeshi, F.M., Shiri, M.E., Faraahi, A.: Using Fuzzy Theory to Information Predicate for Improvement of Cache Operation in Mobile Databases. Future Networks, P104–P107 (2009)
4. Liao, W., Tang, G., Jing, N.: Hybrid indexing moving objects based on velocity distribution. Chinese Journal of Computers 30(4), 661–671 (2007)
5. Liao, W., Jing, N., Zhong, Z.: An efficient prediction technique for range aggregation of moving objects. Journal of Computer Research and Development 44(6), 1015–1021 (2007)

Lifelong Learning in Taiwan, China, Korea and Japan

Su Yen-Ju[1], Chen Shun-Chieh[1], and David W.S. Tai[2]

[1] National Changhua University of Education,
Department of Industrial Education and Technology,
No.2, Shi-Da Road, Changhua City, Taiwan
[2] Hungkuang University, No.34. Chung-Chie Rd.,
Shalu, Taichung County Taiwan
{yenju5329,chensjtw,david60123}@gmail.com

Abstract. Comparing lifelong learning in Taiwan, china, Korea and Japan, according to the EU's lifelong learning quality indicators, Taiwan's overall development of lifelong learning is better than China's but the quality is far lower than that of neighboring countries. Among the 4 areas, Taiwan has a higher rate of participation in lifelong learning in area B and C. Taiwan government should follow the example of neighboring countries, regular public participation and lifelong learning conditions survey conducted as a reference for policy improvement.

Keywords: lifelong learning, PISA, Quality Indicators.

1 Introduction

Human capital signifies a country's true competitiveness. The 21st century is the age of the knowledge economy; it is a century of fierce global competition. Facing the new century, Taiwan's key to success in increasing its national competitiveness, sustainable developing the natural environment, and improving quality of living, lies in education. In particular, the ability to successfully help individuals in their pursuit of free and dignified growth, as well as diverse and orderly social progress, relies on education. A society that invests heavily in education is investing in a beautiful future. Regarding the concept of international lifelong learning, a country's current development and trend of lifelong learning should be emphasized. In Asia, the development of lifelong learning policies of our neighboring countries, such as China, Korea and Japan, especially deserves investigation.

2 Literature Review: Development of Lifelong Learning in Taiwan, China, Korea and Japan

Lifelong learning should not merely be an abstract principle or an ideal but also incorporate concrete policies and actions. According to the education reports released since the 1980s, lifelong learning has become an important policy for implementation for any particular country.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 163–170, 2011.

2.1 Taiwan

From the "Toward the Learning Society White Paper" presented by the Ministry of Education in 1998, it can be seen that Taiwan's government was responding to the global trend towards lifelong learning, which named the year "The Lifelong Learning Year." The Lifelong Learning Act passed in 2002 infused new energy into the promotion of lifelong learning in Taiwan [1]. The 2010 National Education Conference also included lifelong learning on the agenda. A new education white paper will be framed according to the conclusions of the conference and will serve as the policy direction and guidelines on the future development of lifelong learning in Taiwan.

The implementation policies include the following: 1.The supplementary and continuing education system. 2 .Providing learning opportunities for the public by establishing community colleges: currently, there are 101 community colleges and 250,000 students [2]. 3. Accreditation for informal learning: currently there are 544 certified courses (1388 credits) provided by 133 institutes [3]. 4. Providing in-service education opportunities. 5. Promoting learning for senior citizens: the senior population (over 65 years old) has dramatically increased to 2.48 million at the end of 2010, comprising 10.74% of the total population [4]; schools, township offices, senior citizen centers, community service centers, and community development associations started to establish the Learning Resource Centers for Senior Citizens in 2008. It is expected that the goals of establishment will be achieved in 2011 and provide more learning opportunities for local senior citizens [1].

The future development strategies are as follows: 1.Establishing a promotional unit for each level of government and setting a complete structure. 2. Raising the legal rank of lifelong learning and amending the Lifelong Learning Act as soon as possible. 3. Establishing a complete mechanism for achievement certification for lifelong learning. 4. Coordinating and integrating lifelong learning institutes. 5. Providing sufficient promotion and funding for lifelong learning [1].

2.2 China

China has started to develop lifelong learning in recent years due to the increased demand for lifelong learning, adjustments to the industrial structure, economic changes and diverse individual learning needs. In addition, lifelong learning has been greatly developed since the Communist Party of China's Report of the 16th Congress, which included the following: "to build a moderately prosperous society, a lifelong learning society for all by 2020" [5]. Many provinces, starting from Fujian in 2005, have passed laws related to lifelong education [6]. In the 2010 "National Mid- and Long-term Education Reform and Development Outline," China explicitly presented goals and missions such as "constructing a flexible and open lifelong education system," "forming a basic learning society," "building lifelong learning bridges," and "establishing a credit accumulation and transfer system for continuing education, and recognizing and connecting different types of learning achievements" [7].

2.3 Korea

Korea started to actively promote lifelong education in the 1990s. The Education Reform for the New Education System issued by the Presidential Commission on

Educational Reform in 1996 [8] emphasized that a national lifelong learning policy and basic structure should be developed, and that learning opportunities should be open to the public. In particular, a complete system for improving learning sources, support services and certification of learning achievements should be established in order for individuals to learn whenever and wherever it is suitable for their needs and conditions. In 1999, Korea promoted lifelong education by passing the Lifelong Education Act, establishing organizations in charge of lifelong learning and supporting organizations, consolidating the function of lifelong education in colleges, promoting lifelong e-learning, and implementing "credit bank" and "credit recognition" systems, which include setting up a banking system for specialized information and education accounts. Furthermore, the qualification system for instructors was also completed by upgrading social education professionals to lifelong learning instructors [9].

2.4 Japan

Lifelong learning has become a basic state policy of the Japanese government and has long been rooted in their citizens' mind. As early as 1981, the Central Council for Education issued the Report on Lifelong Learning, in which the concept of lifelong integrated education was presented from a broad perspective. The report emphasized that individual lifelong learning should be encouraged in the overall education system. In 1990, Japan issued the Lifelong Learning Promotion Act to build a solid legal foundation for lifelong learning. The Act aimed to provide thorough promotional measures and lifelong learning opportunities [10]. The Japanese government believes that a lifelong learning society and all necessary measures must be established so that "people can choose to learn anytime in their lives and their achievements will be properly assessed." The reasons are as follows: 1. Eliminating the drawbacks of society that overvalue a person's educational background; this is one of the key issues in Japan's current education reform. The aim is to build a society where all kinds of "learning achievements" throughout a person's lifetime can be truly evaluated, regardless of their forms. 2. Reflecting the increased demand for learning as a society matures. 3. As society and the economy changes, lifelong learning becomes even more necessary. People must continually acquire new knowledge and technology in order to respond to information-based lives, globalization and changes in industrial structures [11].

In the network of lifelong learning, community-based lifelong learning centers are established in addition to centers for citizens, especially women. The multifunctional centers offer facilities of all kinds, including: classrooms, information provision, consulting services, entertainments, book loans, and other activities. Besides classrooms, larger facilities also provide learning information or education services, such as consulting with a specialist [10].

3 Comparison

The criteria for comparing the four countries' lifelong learning areas are based on the 15 quality indicators in the 2002 European Report on Quality Indicators of Lifelong Learning [12] presented by the European Union.

The four areas and 15 indicators are as follows Area A: skills, competencies and attitudes, including: literacy, numeracy, new skills for the learning society, learning to learn skills, as well as active citizenship, cultural and social skills; Area B: access and participation, including: access to lifelong learning and participation in lifelong learning; Area C: resources for lifelong learning, including: investment in lifelong learning, educators and learning, as well as information and communications technologies (ICT) in learning; Area D: strategies and systems, including: strategies for lifelong learning, coherence of supply, guidance and counseling, accreditation and certification, and quality assurance. Here we only compare the four countries according to the indicators that have detailed criteria.

3.1 Skills, Competencies and Attitudes in Lifelong Learning

Literacy: Percentage of students per country at proficiency level 1 or below on the PISA reading literacy scale.

Table 1. Differences in reading literacy among Taiwan and neighboring countries in Asia

Country / Level	Taiwan	Japan	Korea	Hong Kong	Shanghai	OECD Ave.
Level 1 and under (%)	15.6	13.6	5.8	8.3	4.1	18.8

Source: [13]

Although Taiwan's overall reading proficiency rate is not especially low, 15.6% of Taiwanese students are at the proficiency level 1 or below, compared to Korea at 5.8%, Hong Kong at 8.3%, Shanghai at 4.1%, and OECD average at 18.8%. It shows that Taiwanese students still have great potential for enhancing their reading proficiency.

Numeracy: Percentage of students per country level 1 or below on the PISA mathematical literacy scale.

Table 2. Differences in mathematical literacy among Taiwan and neighboring countries in Asia

Country / Level	Taiwan	Japan	Korea	Hong Kong	Shanghai	OECD Ave.
Level 1 and under (%)	12.8	12.5	8.1	8.8	4.8	22

Source: [13]

There are at least 87.2% of Taiwanese students above level 1. The OEDC average at level 1 and under is 22% while Korea, Hong Kong and Shanghai are below 10%. It shows that enhancing math proficiency at the lowest level should deserve more attention in our future education policies.

New Skills in the Learning Society: Percentage of students per country below the score of 409 points on the PISA scientific literacy scale.

Table 3. Differences in scientific literacy among Taiwan and neighboring countries in Asia

Country / Score	Taiwan	Japan	Korea	Hong Kong	Shanghai
Score	520	539	528	549	575
Rank	15	6	7	3	1

Source: [13]

There are no significant differences in scientific literacy scores among Taiwan and the neighboring countries; all of them score high.

Table 4. Percentages of science and technology students in higher education in Taiwan and neighboring countries in Asia

Country / Score	Taiwan	Japan	Korea[1]	China
Science and technology students in higher education: colleges (%)	34.5	18.6	38.7	41.7
Science and technology students in higher education: MA & PhD (%)	42.2	38.9	23.0	48.3

Source: [1, 14] [1.] The data for Korea includes science, technology and agriculture.

The percentage of Taiwanese and Chinese students pursuing Masters/ Doctoral degrees in science and technology are over 40%, showing that both countries accept new technologies at a relatively high level.

Learning-to-Learn Skills: Percentage of students per country in the lower 25 % of overall performance on the PISA "elaboration strategies" index.

Table 5. Differences in elaboration strategies among Taiwan and neighboring countries in Asia

Country / Score	Taiwan	Japan	Korea	Hong Kong	Shanghai	OECD Ave.
Index	460	494	512	527	544	489

Source: [15]

In the lower 25% of overall performance, Taiwanese students obviously score lower than the neighboring countries in Asia and OECD average, showing that Taiwan needs to enhance teaching skills in reading and information acquisition skills. It also shows that Taiwan's competency in learning-to-learn skills, one of the important indicators for lifelong learning, is obviously lower than other countries'.

Active citizenship, cultural and social skills: Civic knowledge and interpretative skills, civic knowledge, civic engagement and civic attitudes across countries.

Table 6. Differences in civic awareness among Taiwan and neighboring countries in Asia

country / score	Taiwan	Japan	Korea	Hong Kong	International Ave.
Index	559	--	565	554	500
Rank	4	--	3	5	

Source: [16]

3.2 Participation in Lifelong Learning

Participation in education and training of those ages 25 to 64.

Table 7. Comparison of dropout rates (ages 18-25) among Taiwan and neighboring countries

country	Taiwan	Japan	Korea	China	Hong Kong	OECD Ave.
Percentage (%)	0.03	1.48	0.6	2.08	0.181	13

Source: [1, 17]

3.3 Resources for Lifelong Learning

Investment in Lifelong learning: Total public expenditure on education as a percentage of GDP.

Table 8. Comparison of percentages of GDP on education among Taiwan and neighboring countries

Country	Taiwan	Japan	Korea	China	OECD Ave.
Percentages of GDP on education (%)	4.9	3.4	4.2	3.3	5.2

Source: [1, 18]

Educators and Learning: Percentage of teachers having received education and training during the previous four weeks.

Table 9. Comparison of the HDI index among Taiwan and neighboring countries in Asia

Items / Country	HDI (2010)	Rank	Average years of education	Rank	Expected years of education	Rank
Taiwan	0.868	18	11	22	16.1	17
Japan	0.884	11	11.5	17	15.1	36
China	0.663	90	7.5	94	11.4	122
Korea	0.877	12	11.6	12	16.8	9

Source: [19]

The Human Development Index (HDI) was created by the United Nations Development Program (UNDP) in 1990, considering that although “income growth” is indispensable in human development, it’s not the only focus. In 2010, UNDP updated the measurement index to reflect changes over time, combining “life expectancy at

birth," "average years of education," "expected years of education," and "average GNI per capita by purchasing power parity" into one index to more accurately reflect a country's development in the areas of health, education and economy. There are no significant differences in the overall index performance among Taiwan, Japan and Korea; China's HDI performance, on the other hand, ranks low among 127 countries surveyed. As for the education index, Korea's average years of education ranks first among the four countries.

ICT in Learning: Percentage of households who have internet access at home.

Table 10. Comparison of the IT training and education index among Taiwan and neighboring countries in Asia

Country	Taiwan	Japan	China	Korea
Index	63.68	20.08	33.24	44.21
Rank	14 (Ms 3/7)	64 (Ls 6/6)	53 (Ls 5/6)	41 (Ms 7/7)

Source: [20]

3.4 Strategies in Lifelong Learning

Member states' positions on developing lifelong learning strategies.

Table 11. Comparison of employment rates by education levels (ages 25-64) among Taiwan and neighboring countries in Asia

Country / Education	Taiwan[1]	Japan[2]	China	Korea[2]	OECD Ave.[2]
Below upper secondary	49.5	67.6	--	66.1	58.7
Upper secondary and post-secondary non-tertiary	62.34	74.4	--	70.7	76.1
Tertiary education	69.04	79.4	--	77.1	85

Source: [18, 21] [1.] 2010 "Human Resources Survey" by the Directorate-General of Budget, Accounting and Statistics, Executive Yuan, R.O.C. [2.] 2001 OECD data

4 Conclusion

If the four countries are compared according to the EU's lifelong learning quality indicators, Taiwan's overall development of lifelong learning is faster than China's but the quality is far lower than that of neighboring countries. Among the five indicators in area A, all but "citizen awareness" is lower than other countries; in area B, Taiwan has a higher rate of participation in lifelong learning; in area C, Taiwan leads the other three countries in investment in education, educator and learning, and ICT in learning; in area D, only the first indicator (strategies for lifelong learning) is defined and compared. The employment rates of the neighboring countries are lower than the OECD average presumably because the economic conditions affect employment at various education levels; therefore, all of the countries need to improve their policies for lifelong learning.

The Taiwan government can't really understand the concept of the participation in lifelong learning, but can not set the policy that corresponds to the future to improve the

situation. To neighboring countries as an example, Japan, Korea, regular people across the country for lifelong learning participation and level of investigation, to understand the needs, satisfaction, and participation rate, and with the population of background information to conduct cross-analysis to further understand the all age, sex, occupation types of cases involved, and compile this information, write reports, set as a basis for policy. Taiwan government should follow the example of neighboring countries, regular public participation and lifelong learning conditions survey conducted as a reference for policy improvement.

References

1. Ministry of Education, MOE, Taiwan, http://www.edu.tw/
2. National Association for the Promotion of Community Universities, http://www.napcu.org.tw/website/category/data/statistics
3. Non-formal Education Programs Accredtiton Center, Taiwan, http://140.122.96.1/p8.htm
4. Department of Statistics, MOI, Taiwan, http://www.moi.gov.tw/stat/
5. Han, M.: China's progress in lifelong learning policies. In: International Conference on Lifelong Learning Policies Report (2009)
6. Ministry of Education, MOU, China, http://www.moe.edu.cn/
7. The Chinese adult education association, http://www.caea.org.cn/caea_content.asp?id=352
8. Ministry of Education, Science and Technology, http://english.mest.go.kr/web/1699/site/contents/en/en_0209.jsp
9. Lee, J.-B.: Vision and Strategy for Developing Lifelong Learning in Korea. In: KEDI (2004), http://eng.kedi.re.kr/khome/eng/archive/report/viewReports.do
10. Huang, W.-S.: Comparative Education, Hedu, Taipei, pp. 214–217 (2011)
11. Ministry of Education, Culture, Sports, Science and Technology. Lifelong Learning Policy (2010), http://www.mext.go.jp/english/
12. European Report on Quality Indicators of Lifelong Learning. In: European Commission, Brussels (2002)
13. PISA, Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science, vol. I. OECD Publishing, OECD (2009)
14. International Comparison of Education Indicators. In: Mext (2011)
15. PISA 2009 Results: Learning to Learn: Student Engagement, Strategies and Practices PISA, vol. III. OECD Publishing, OECD (2010)
16. Schulz, W.: ICCS 2009, International Association for the Evaluation of Educational Achievement(IEA), Amsterdam (2010)
17. Ho, J.-T., Wu, J.-C., Wang, S.-C., Yang, S.-L., Cheng, C.-K., Tsai, P.-S.: The Current Conditions of Dropouts and Guidance Measures in Various Countries—Taiwan's Guidance Policies on Returning to Schools for Dropouts in the Past and Future, p. 4. Ministry of Education Students Affairs Committee, Taipei (2003)
18. Education at a Glance 2010. OECD Indicators, OECD (2010)
19. Directorate General of Budget, Accounting and Statistics, Executive Yuan, R.O.C. Statistical Bulletin conditions (2011), http://www.dgbas.gov.tw/public/Data/11715541971.pdf
20. IPS National Competitiveness Research 2010-2011, IPS and IPS-NaC, Seoul (2011)
21. Directorate General of Budget, Accounting and Statistics, Executive Yuan, R.O.C. Manpower Utilization Survey (2010), http://www.dgbas.gov.tw/public/Attachment/01281461071.pdf

Design of the Portable Gloves for Sign Language Recognition*

Xu shijian, Xu jie**, Tang yijun, and Shen likai

School of Information and electrical Engineering,
Zhejiang University City College
Hangzhou 310005. China
Xushijian1226@163.com

Abstract. In order to improve the gesture communication ability of the obstacle of hearing people, a portable gesture recognition data-glove is designed, which uses 14 groups of three-axis acceleration sensor- totally and 42 acceleration channels, and Bluetooth to transfer the data to the embedded system. Obtaining and identifying the effective parameters through the real-time analysis of the 42 groups of variables. Based on the identification, gesture converts into semantic conversion, and finally semantic conversion converts into voice, thus to realize the gesture transforms into voice in real-time.

Keywords: gesture recognition; data collection; portable; data-glove; Embedded System.

1 Introduction

Sign language used by the obstacle of hearing people is a special language communicated by action and visual, which is a stable expression system constituted by hand gestures and expressions. The purpose of sign language recognition is get the meanings of sign language data obtained by computer data acquisition device and translate into voice, which uses the pattern recognition algorithm and consults of the context. [2]

At present, whether based on visual or data gloves, the terminal of sign language recognition research is computer, which is not convenient to carry. Therefore, based on the previous studies, this design does further research on sign language recognition technology and make it portable. Portable hand gesture recognition system can identify some basic gestures and display the meanings on the PDA terminal. This data gloves' design greatly improves the applicability and the recognition accuracy, enhances the users' experience degree, which has great marketing values.

* This work was supported by Students Scientific and Technological Activities ("XinMiao" Technicians Program) in Zhejiang Province 2011 and funded by the Zhejiang Province Education Science Research & planning project(SCG107).

** Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 171–177, 2011.

2 Comparative Analysis of Gesture Recognition Technology

The key premise that gesture used in human-computer interaction is the recognition and explanation of the gesture inputs through computer, people take different kinds of measures to recognize gestures now: [4].

Based on the mouse and pen, the disadvantage is that it can only recognize the entirety motion of hands but lost the figners' movement. The advantage is suit for most desktop systems, because it carries out by software algorithm. To more explication, only when using the mouse cursors' or the pen's movement or the changes of direction to convey the information, can regard the mouses or pens as the tools of gesture expression. This kind of technique can be used in text collating and so on.

Based on the data gloves, it can determine the fingers' posture and gestures, but otherwise relatively expensive, and it is inconvenience sometimes, such as perspire.

Based on computer vision, which uses the camera to import gestures. Its advantage is noninterference to users. Many researchers devote themselves to this promise work now.

However it has many technical difficulties, the current computers are under qualified for the task of gesture recognition and under standing.

2.1 Summary of Gesture Recognition Technology

From the gesture input devices, the current research of gesture recognition system is mainly devided into: data gloves-based and vision-based.

The main gesture recognition technology obtained now is: [3].

The Template matching technique is a simplest recognition technique used in static gesture recognition mostly. It uses the sensor input of raw data to match with the template stored in advance, and completes the recognition task by measuring the similarity between the two. Such as: Template by calculating the correlation coefficient to predict the match.

Neural network technology is a relatively new recognition technology, which has self-organizing, self-learning ability and has the characteristics of distribution .what's more, it can be effective in suppressing noise, deal with the incomplete model and has the capacity of mode promotion.

Statistical analysis technique is a classification method based on probability through gathering statistics of the sample's feature vector to determine the sorter. In the pattern recognition, Generally uses the Bayesian maximum likelihood theory to define the classification function. But the disadvantage is asking people to extract the feature vector from the original, nor identify the original data directly.

2.2 The Problem of Gesture Recognition Technology Based on Computer Vision

The operational process of gesture recognition system based on vision obtains video data stream by one or more cameras. And then, the system accorded the interaction modeling to check whether there is gesture in the data stream. If got, the system divides the gesture from the video signal, then, chooses gesture model to analyse. The analyzing process consists of feature detection and model parameter recognition, classifies gesture and describes according to the model parameters if needed. At last, the system based on the descriptions to drive the application.

But the most researches stress on static hand gesture recognition, it has some technology difficulties:[5]

The difficulty of gesture target detection.

Difficult to recognize gesture goals: the real time intercept of goal is to intercept the goal out from the image stream in the complex background, it is one of the main subject about machine vision research. Nowadays, there are many mature and easy-achieved technologies aimd at self-motion vision system.

The gesture recognition is to explain a higher meaning based on the poses and changes of people's hands, and its key technology is to pick up the characteristics with geometrical invariability. The gesture has these characteristics:

There are large differences among one kind of gesture because the hand has resilience.

The main part of gesture recognition is finger, so the part of palm characteristics is redundant information.

It is hard to orientate the position of hand in the three-dimensional space, so the projection is important because the picture that the computer gets is from three-dimension to two-climension.

The surface of hand is slick, so it is easy to produce shadow.

2.3 The Advantage of Data Glove-Based Gesture Recognition Technology

Data gloves are a kind of Interactive equipment which widely used in virtual reality technology. Actually, traditional interaction devices also can regards as gesture input devices, such as a mouse or pen.

Glove-based Gesture Input has the following advantages:

Small amount of input data, high speed, it can get 3-d information on space and figuers' movement.

Many kinds of gestures information can be identified in real time.

Currently, the method of gesture recognition based on data glove is the neural networks mainly. Neural network is suitable for the train in the fast and interactive way, because it can be used as static and dynamic input, without having to define the transfer characteristic in a mode of analyze.

Gesture recognition program can adapt to different users by adjusting the network's connection weights according to users' Personal circumstances [1].

The method of the gesture language recognition based on ANN / HMM sign is compounding application with data glove CyherGlove, owning 18 models of sensor. It turned out to be effective and feasible in the gesture recognition system.

3 The Design of Portable Sign Language Recognition Glove

3.1 The Framework of Sign Language Recognition Glove Design

The design is divided into two parts, hardware and software. Hardware needs the acceleration sensor technology, Bluetooth communication technology, and single chip communication control technology for research .Software needs the WinCE embedded system programming, mathematical modeling and so on for research.

Figure 1 is a block diagram of the system.

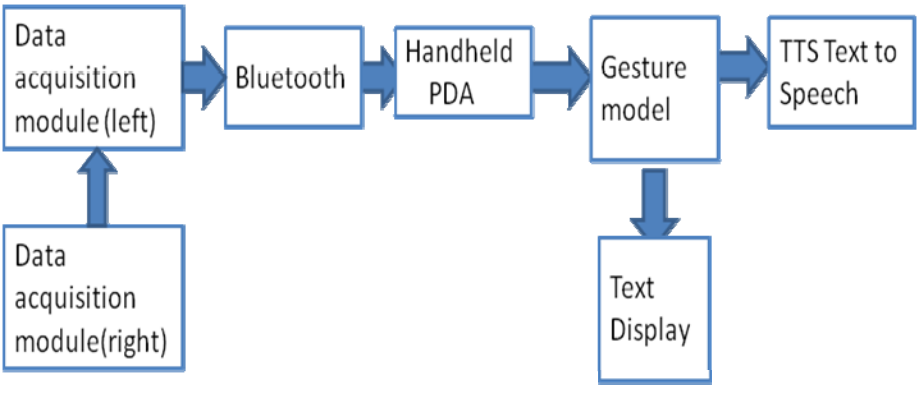


Fig. 1.1. General block diagram of the system

3.1.1 System Hardware Structure

The data acquisition module: Uses three-dimensional acceleration sensor to capture the gesture datas, and Transfers them to the single-chip microcomputer.

Through the high-precision AD transfer module of the single-chip microcomputer, the acquisition of analog signals are converted into digital signals, then the datas will be transmitted to the PDA via Bluetooth.

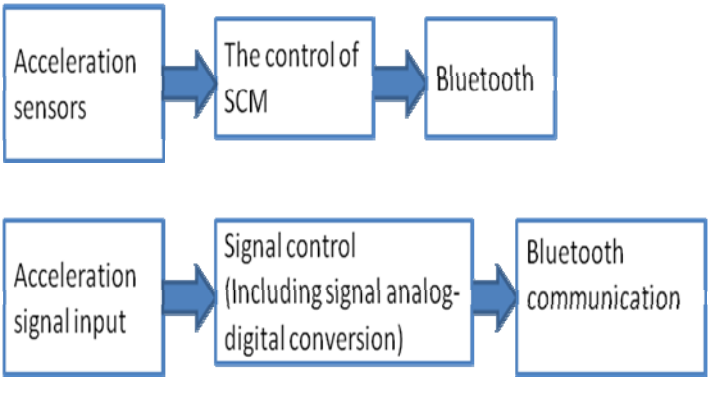


Fig. 1.2. Block diagram of data acquisition hardware module

3.1.2 System Software Architecture

The software is designed as three layers, the presentation layer, data processing layer and persistence layer.

The presentation layer: orients the users used for text and speech output, which is the interaction layer between software and user .And it is used for text and speech output.

The data layer: It is used to do the characteristic operations and pattern recognition on the collected data.

The persistence layer: It encapsulates the data access details, which is the SQLite database access layer.

SQLite is a light-weight relational database management system that observes ACID, design goal is embedded. It is widely used in the multitech for it only needs husbands of K memory .the SQLite supports many mainstream operating systems such as Windows/Linux/Unix, meanwhile it can combine with lots of programming languages like Tcl, PHP, Java and so on, also ODBC joggle .Compared with Mysql, PostgreSQL, the two famous original database management systems, the SQLite is speeder.

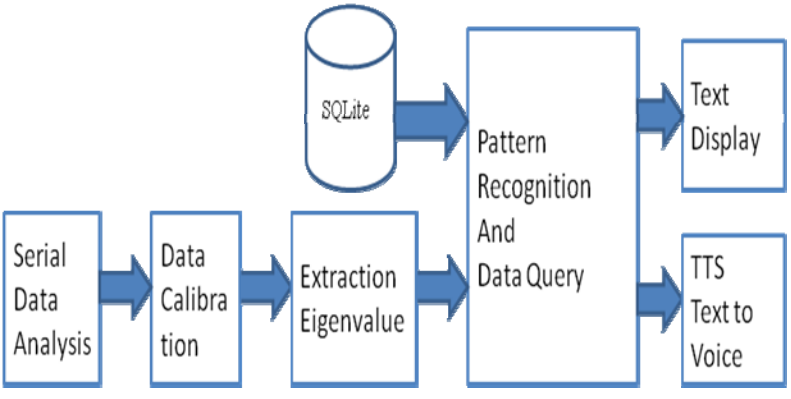


Fig. 1.3.

3.2 System Workflow

3.2.1 Data Collection Process

By changing the hand posture 14 groups of 42-channel acceleration sensor loaded on the glove changes, And the single-chip microcomputer collect the sensors' analog signals constantly. Through the high-precision AD transform module, the analog signals are converted into digital signals. Which transmiss to the handheld device via Bluetooth. Finally, the data collection of gesture is completed.

3.2.2 Data Recognition

The program is begin with initializing the microcontroller and acceleration sensors. And then SCM start to receive sensor data to start the AD conversion module, which make AD change of the analog signal.The SCM uses a I^2C- bus communication in each other, and the collected datas are sent to the host machine, which sends the changed data to the PDA via Bluetooth. The PDA takes a eigenvalue extraction after receiving the data, and then match the searching between Characteristic values and the model. If found, shows the result, otherwise continues to the next data processing.

The system process flow diagram as shown in Figure 1.4:

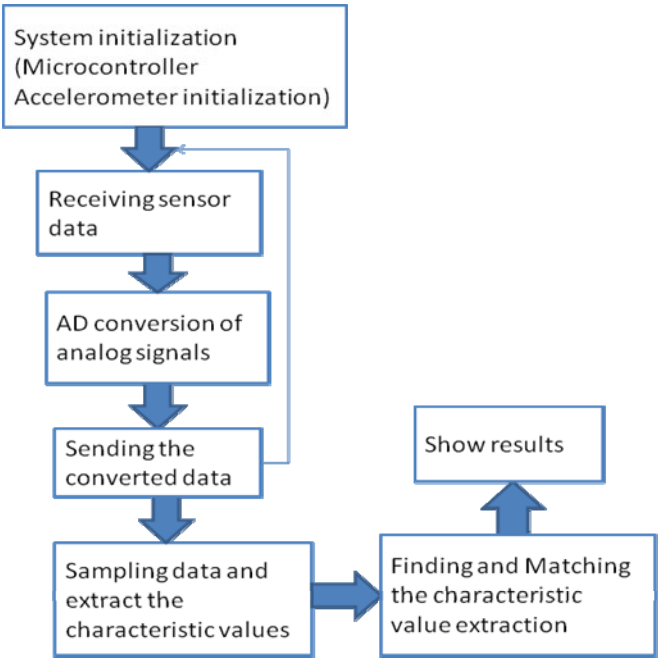


Fig. 1.4. Program flowchart

4 Technical Characteristics and Innovation System

The single I^2C bus in the project is mainly used the Multi-microcontroller data communication by the mode of one main more sub, The five pieces of data collected by machine are sent to a host through the I^2C bus communication, then the host are the data to the handheld devices via Bluetooth, Data can be viewed through the serial port. Its advantages are:[6]

1) General computing applications in system design: The system uses portable equipment to ensure the system can provide services to users anytime, anywhere.

2) User-friendly applications designed in handheld devices: Single chips and three-dimensional acceleration sensors will be packaged with a personalized design to make gloves。

3) Electronic devices used in the systems have low cost, the circuit is simpe and the development cost is low.

4) Bluetooth technology, makes the data collection gloves be separated from the hand-held devices which are more flexibility.

5) The terminals of traditional gesture recognition technology are on the computer equipment,, which is inconvenient to carry. But this project make a breakthrough on this point, the terminal is a handheld device, which is convenient to carry, and the data can be transmitted between gloves and hand-held devices by Bluetooth. So the hand-held device can move flexible in the range of Bluetooth transmission.

6) Multi-platform portability: the system's portability is taken into account in the early design, So that it can apply for all hardware platforms of the Windows CE operating system. And it also supports screen operation. The supported platforms: Microsoft Smart Phone 2003; Microsoft Pocket PC 2003; Windows Mobile 5.0 Pocket PC; Windows Mobile 5.0 Smart Phone and so on.

5 Conclusion

Based on the data glove sign language recognition is one of the future development trend. This article outlines various aspects of sign language recognition, and discusses the technology to achieve this recognition system.

The nearest closed to sign recognition is gesture recognition which is the research hotspot in the human-computer interaction after voice recognition, and has widely foreground. For instance: it is used in interactive of virtual environment, hand gesture recognition, the grabbing of robot's hand. Sign language as a action language, not only applies to among the obstacle of hearing people, with normal persons, but also can be widely used among the normal person. So the research of gesture recognition based on the data glove has extensive social significance and actual application prospect in the human-computer interaction.

The study of the system helps to improve the deaf-mute's living, learning and working conditions, provides them with better service, meanwhile it contributes to computer-aided sign language instruction, bilingual education television, virtual human research and so on.

References

1. Li, Q.-S., Fang, Z.-G.: Gesture recognition technology and the application in human computer interaction. Human Engineering Education 8(1) (2002)
2. Liu, H.-K., Qi, G.-X., Chen, C.-H.: The research and application of new gesture language recognition. Machine Tools and Hydraulic 35(7) (2007)
3. Heap, T., Hogg, D.: Wormholes in shape spacetracking through discontinuous changes in shape. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 344–349 (1998)
4. Imagawa, K., Taniguchi, R., Aftra, D., et al.: Appearanc gased recognition of hand shapes for sign language in low relutionimage. In: Proceedings of the 4th Asian Conference on Computer Vision, pp. 943–948. IEEE, Los Alamitos (2000)
5. Hu, Y.-S.: The summarize of gesture recognition. China Science And Technology Information 2 (2009)
6. Lu, K., Li, X.-J.: The research description of gesture recognition. Journal of Xi'an University of Arts and Science (September 2006)

FEA Simulations of Temperature Changes of Castings When Solidified and Cooled in Mould*

Xinping Hu[1,2] and Ling Xie[1,2]

[1] School of Mechanical Engineering, Shandong Polytechnic University, Jinan, Shandong, 250353, China
[2] Shandong Provincial Key Laboratory of Advanced Machine-making of Light Industry and Monitoring Technology, Jinan, Shandong, 250353, China
{huxxpp,xieling2045}@163.com

Abstract. The temperature distribution of casting in mould is a key factor affecting the quality of castings. In this paper, we try to get the Law of temperature distribution in the casting process by numerical simulation with software ANSYS. The whole process can be summarized as follows: creating mathematical model, setting the initial and boundary conditions of the mathematical model, analyzing the solidification time of the casting, making out the temperature distribution in the whole process, and evaluating the simulation results. The results indicated that there existed self tempering effect in the transitional section of the casting, and the temperature distribution of casting was close to actual production.

Keywords: ANSYS; Casting Mould; Temperature; Self Tempering Effect.

1 Introduction

The quality of castings is closely related to the distribution of the surface temperature of mould cavity. So the distribution of the surface temperature of mould cavity has a direct affect on the quality of the casting and the life of the mould. In actual production, its temperature field and thermal stress field are being in an unsteady state, it means that the property parameter changes with time. However, once the design and manufacturing of the dies are completed, it's difficult to make a big change even if some errors are found. Thus, using numerical technology when we design a die and predicting the temperature distribution of various parts of the mould and its variation is very important to improve the casting quality, prolong the service life of mould, reduce costs and speed up the upgrading of products [1].

The temperature field of the mould makes an effect on the conditions of filling and cooling of the castings, and then influences the quality of its forming quality and internal organization. On the other hand, the dies endure sudden temperature change repeatedly and the capability level to thermal stress is determined by the temperature

* **Foundation item:** Project(J09LD11) supported by Shandong Province Higher Educational Science and Technology Program; Project (BS2009ZZ010) supported by Shandong Province Outstanding Research Award Fund for Young Scientists.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 178–183, 2011.

field of the mould. It is well known that the major reasons for die failure are thermal fatigue, wear and deformation when aluminum alloy casting is produced with die mould.

Temperature distribution of mould in production process can be achieved by finite element analysis, effectively. In this paper, we will use ANSYS to analyze the temperature distribution during different casting conditions and hope to increase the die life and improve casting properties by means of improving the distribution of temperature field.

2 Mathematical Model in Casting Process

2.1 Mathematical Model

In casting process, after the liquid metal poured into the mould, the process of its solidification and cooling in the cavity is a process of heat emission through the casting mould. When the molten metal is injected into the mould cavity, it releases large amounts of heat in a very short period of time, raising the temperature of the mould in a flash. At the same time, the die absorbs some of the heat through conduction, convection, radiation, from surface coatings and cooling water, so the temperature of the die drops. After a period of time, the temperature of the mould achieves a relative balance point.

From pouring to solidification and cooling, the heat transfer system in casting mould depends on three ways such as the radiation heat of the high-temperature metal, the convective heat transfer between liquid metal and the casting (including the convection heat transfer between mould surface and atmosphere), the heat transfer from the metal to the mould and so on. When the liquid metal filled in the cavity completely, assuming that there is no convection in the liquid metal when the solidification is underway, the solidification process is basically seen as an unstable process of heat conduction [2].

According to the characteristics of casting, the three-dimensional unsteady temperature field is adopted here, the control equation is shown below,

$$\rho C_p \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}(\lambda \frac{\partial T}{\partial x}) + \frac{\partial}{\partial y}(\lambda \frac{\partial T}{\partial y}) + \frac{\partial}{\partial z}(\lambda \frac{\partial T}{\partial z}) + Q \tag{1}$$

In the equation above, ρ- Density, C_p- Specific heat, t-Time, T- Temperature, λ-Thermal conductivity, Q-The heat source per unit volume.

Enthalpy method was used to deal with the latent heat of crystallization. A mathematical model was created in our study which was shown in Fig.1.

In Fig. 1 (a), the models which consisted of two kinds of materials were shown with grids. The inner zone was Al-Si casting, and the outer zone was copper made permanent die. The wall thickness of casting changed from 20mm to 2mm gradually, the height of the casting here was 30mm, and the dimensions of the die were shown in Fig.1 (b). In the model shown in Fig.1 (a), there were total 950 nodes, in which 258 nodes belonged to Al-Si casting, and the other 692 belonged to the permanent die. In Fig. 1 (a), three nodes marked A, B and C were on behalf of three different zones with different wall thickness respectively. A was on behalf of the thick wall zone and its

node number was 101, C the thin wall zone and its node number was 171, B the transitional zone and its node number was 224. The reason why A, B, C were chosen in the center of the casting was to avoid the chill effect.

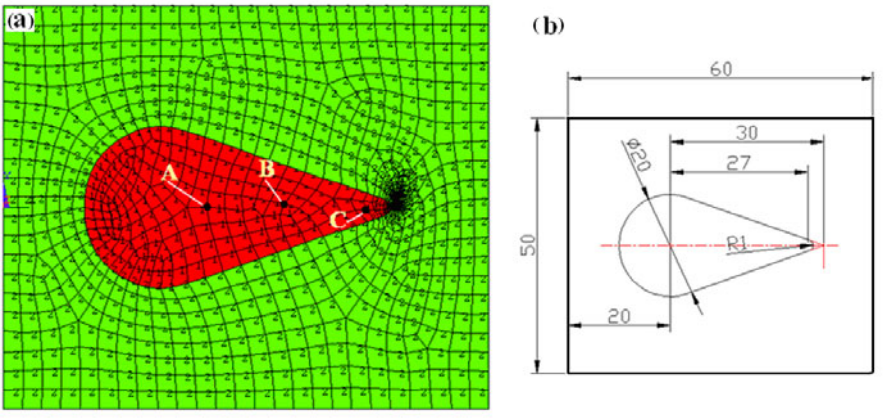


Fig. 1. Diagram of FEM temperature analyses modeling with meshing grid: (a) analyses modeling with meshing grid; and (b) modeling with dimension in mm

2.2 Determination of Conditions for Solving

In the numerical simulation technology, the boundary and initial conditions should be determined according to practical problems after the mathematical model was created. The boundary and initial conditions are the main factors affecting the calculation accuracy.

In this analysis, the mould was made of copper and the casting was Al12Si alloy. Pouring temperature is 690℃ and the preheating temperature of the mould is 200℃. For the small castings, Pouring time is very short. High-temperature liquid aluminum alloy contacted the low temperature mould suddenly. The boundary temperature wouldn't change too much. So the boundary temperature was the preheating temperature 200℃, the initial conditions was to take the pouring temperature 690℃ [3].

Table 1. Parameters of material and thermal conditions

Parameter(s)	Al-Si alloy				Copper
Specific heat : c	837 (J/kg/K)				307 (J/kg/K)
Coef. of thermal cond.: λ (W/M$^{/0}$C)	155 20℃	168 577℃	77 600℃	83 700℃	391
Density: ρ (Kg/m^3)	2650				8940
Die temperature	-				200℃
Pouring temperature	690℃				-
Eutectic temperature	577℃				-
Compositions: wt%	Si:12% ; Al: Balance				-

Casting and mould were closely linked, only heat transfer occurred between them. In the casting process, the surface temperature of the mould was not high, only the convective heat transfer with the air was considered, while the radiation to air was not considered. In this study, the environmental condition was supposed that the temperature of air outside the permanent die was 200℃, and the coefficient of heat transferring was $3W/m^2/℃$. In the FEM analyses, the parameters of material and thermal conditions were shown in Table 1.

3 Finite Element Simulation and Results

3.1 Brief Description of the Steps in Analysis

ANSYS is one of engineering simulation software which is based on finite element method. Its simulation process is taking the continuous medium as lots of discrete units which connected to each other by nodes, forming a whole or aggregate, and using it to create a physical model. Then, create mathematical geometric model and finite element analysis model based on the physical model [4].

(1) Enter thermal analysis environment of ANSYS, and define the thermal analysis unit type as PLANE55. Make transient analysis to the temperature change of nodes A, B and C in the process of cooling and solidification with this thermal analysis unit.

(2) Set physical parameters of materials according to the data shown in Table 1, and import attribute file of the material.

(3) Create analysis model and make meshing grid.

Finite element analysis meshing model of temperature field and its dimension were shown in Fig.1.

(4) Load the model.

Set the initial temperature field according to the parameters of material and thermal conditions shown in Table 1. And define the convectional boundary and the steady-state thermal boundary.

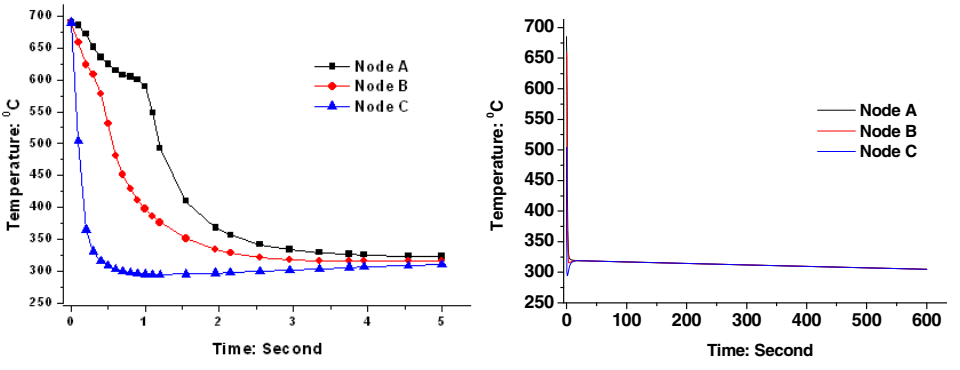
(5) Set the time and the parameter of time step.

Setting the parameter of time step can make ANSYS get to acceptable convergence accuracy in a short period of time. Castings of different sizes have different cooling time in metal founding. Small metal castings are between the 10s ~ 60s, and large sand castings can get up to 20h. All the time steps we set in this analysis are 0.01s.

(6) Set the output control and solve it. Solve it by using transient thermal analysis. Record the results every five time steps in order to study.

We set two periods after liquid Al-Si alloy being poured into the die, they were 5s, 600s respectively. The results were shown in Figure 2(a) (b). In Fig.2 (a), there existed a platform in the temperature graph because of the emitting of latent heat of solidification for node A and node B. But for node C, the latent heat was too small to display a platform in the temperature graph. It could be seen in Fig. 2 (b) that the temperature of node C dropped fast, and then ascended contrarily, then dropped slowly. In the case studied here, the lowest temperature was about 295℃, and then the temperature ascended for a period of time rather than dropped continuously, so this

stage was called self tempering. And the influences on the properties of casting were called self tempering effect. But in general, the temperature decreased as time went by, as shown in Figure 2 (b) below.



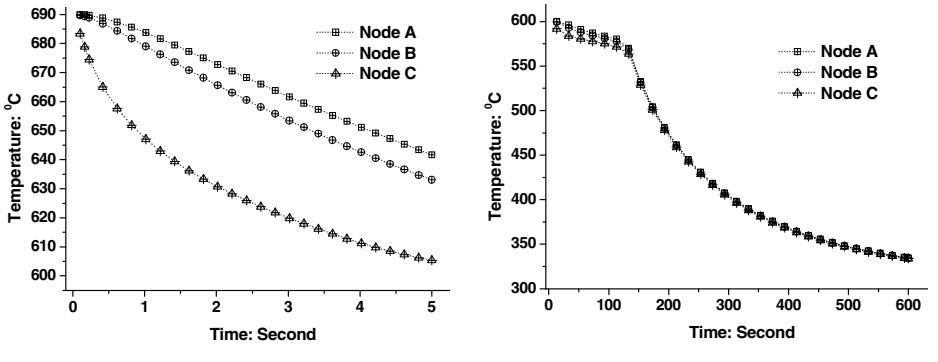
(a) Poured within 5 seconds (b) Poured within 600 seconds

Fig. 2. Temperature changes in different positions cooled in copper-made die

3.2 Simulation Results of Sand Casting

If the dry sand mould was adopted, rather than copper die, the temperature changes of the same position of the Al-Si casting in the same modeling were shown in Fig. 3. The parameter of Al-Si alloy adopted was the same as analyses above, and the parameters of the sand mould were that: density 1580Kg/m3, specific heat 1130J/kg/K, coefficient of thermal conductivity 0.61 W/M/℃ [5].

Like the case used copper die, the temperature graphs were plotted with time duration 5 and 600 seconds after liquid Al-Si alloy poured into the mould. In Fig. 3 (a), there didn't appear a platform and self tempering effect. For all the temperature graphs of node A, B and C, the temperature dropped continuously.



(a) Poured within 5 seconds (b) Poured within 600 seconds

Fig. 3. Temperature changes in different positions when cooled in dry sand mould

From the above analyses, it could be easy to conclude that the mould temperature distribution has its own characteristics for different casting process. When metal casting was adopted, the temperature dropped rapidly and there existed because of the high cooling rate. In first 3 seconds, the temperature of different parts is obvious different. After 5 seconds, the differences became smaller and smaller. In the later period, it's unable to see the difference. While in sand casting, the temperature dropped more slowly [6]. After 5 seconds of solidification, the temperature of different parts still takes on obvious differences. This is mainly because of the high heat transfer capability and quick heat dissipation of metal mould. These results are similar to the actual casting process. If we integrate the characteristics of mould-casting temperature field system, some temperature information on special section of parts could be got and some suggestions could be provided on enhancing casting structure and optimizing parameters of casting process.

4 Conclusions

(1) When the casting solidified and cooled in the mould, in the thin-walled part of the casting or near it, self tempering effect is particularly evident. When metal casting was adopted, different parts of the temperature difference is very small. While in sand casting, the temperature dropped more slowly and the temperature of different parts still have obvious differences.
(2) Temperature field of casting during solidification and cooling can be analyzed by finite element method through establishing mathematical model and setting technological conditions. The results of the analysis were found to be similar to the actual production.

References

1. Zhang, L.Y., Jiang, Y.H., Ma, Z., Shan, S.F., Jia, Y.Z., Fan, C.Z., Wang, W.K.: Effect of cooling rate on solidified microstructure and mechanical properties of aluminium-A356 alloy. Journal of Materials Processing Technology 207, 107–111 (2008)
2. Hu, X.P., Zhao, G.Q., Wang, W.M.: Solidifying pressure and microstructure of AlSi10Cu3 in die sleeve in high pressure die casting. International Journal of Cast Metal Research 23, 289–295 (2010)
3. Hu, X.P., Xie, L.: Studies on the digital expression of Al-Si alloy microstructure and its applications in properties analysis. In: 2010 International Conference on Digital Manufacturing & Automation, ICDMA 2010, vol. 2, pp. 427–431 (2010)
4. Shukla, P.P., Lawrence, J.: Analysis of temperature distribution during fibre laser surface treatment of a zirconia engineering ceramic. In: Proceedings of the 36th International MATADOR Conference, vol. 16, pp. 539–542 (2010)
5. Liu, Z., Mao, W.M., Zhao, Z.D.: Effect of local chilling on morphology of primary α(Al) in semi-solid A356 alloy. Transactions of Nonferrous Metals Society of China 18, 573–579 (2008)
6. Basavakumar, K.G., Mukunda, P.G., Chakraborty, M.: Influence of grain refinement and modification on microstructure and mechanical properties of Al-7Si and Al-7Si-2.5Cu cast alloys. Materials Characterization 59, 283–289 (2008)

The Security System Analysis of Teaching and Researching Management System Based on Audit Strategy

Yuanqiang Wang, Zhihui Wang, Shangfu Hao, and Baili Sun

School of Information Science and Engineering
Hebei North University,
Hebei, Zhangjiakou 075000, P.R. China
{Hbnuwyq,iowzh}@126.com, zjkcxh5656@sina.com,
skutu@163.com

Abstract. By analysis information system security situation and research the requirement of security of Teaching and Researching Management System which was implemented in our projects, a new security system strategy was presented based on audit log files combing with the RBAC access control technology. At the same time, we redesigned and restructured the RBAC model to improve security procedures of our project. With using the new security audit model which provided by us, the security, reliability, forward looking of teaching and researching management system was improved.

Keywords: teaching and researching management system; log file; audit analysis; security model.

With communication technology and computer technology are being applied widely in all kind of fields of society, we find that more and more classified data were stored in different brand disks or the other storage devices which were connected to network. Under stimulating of special interests and objective, at the same time, illegal invasion events were discovered repeatedly, the security problem of information is becoming more serious than ever before. Audit mechanism as the control strategy is the last security barrier of information system which was structured on the security requirements of information system, which formed with the procedure of analyzed, designed and developed. The importance of audit mechanism has been founded clearly in data disaster recovery, system loss evaluation and attack investigation and evidence collection when the disaster really happened.

1 Security Situation of Information System

Today with the rapid development of computer technology, storage technology and network technology, the inevitable outcome from informational construction, information system cannot exists if it nonsupport of operation system, network and database alone. From 1987, Hao Albert-Zinn, the first prisoner for invasion information system was reported on newspaper, to 2011, the number of economic

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 184–188, 2011.

losses which caused by illegal invasion and virus is astronomical. Broadly speaking, the main threats the present information system is facing come from three aspects: the software design flaws, TCP/IP protocol cluster design flaws and artificial operation errors [1]. But the common designer of program was not paid more attention to the security of operation system design and database design, they more concerned with the security problem on the information system which uses by enterprise. GB/T20008-2005, the newest evaluation criteria published by Chinese government referenced the Trusted Computer Standards Evaluation Criteria (TCSEC) which was published by American in 1983. Although it's clearly stipulated by TCSEC that classes above C2 must be assorted with audit system, yet C2 class security still belongs to lower level system, that's why most present systems with their own audition mechanism fail to achieve B class requirement [2].

2 Paper Preparation

The Network Teaching & Researching Management System is a service platform on network which developed for the teaching & researching requirements of county/district level education bureau. The System is a typical C/S web application based on MVC framework, and was developed with the basic environment of Myeclipse+Java+SQL2005.The mainly function is to provide services such as teaching management, teaching & researching activities, video conference, examination room arrangement, score management & analysis, and teacher's growing records. Tow major entities of the System are teaching & researching office and school. Because of the unbalanced distribution of middle and primary schools in this district, together with the unpredictability of users' operation, the district Teaching & Researching Management System, which is running under network environment, undertakes serious operation risks. According to the China's Internet Safety Reports (2010~2011) [3], it's believed that 93.2% of viruses and Trojans are spread directly through internet, in this case, complete defensive measures to ensure the system safety and data integrity under network operation environment are required.

The investigation on management system of Teaching and Researching shows that in order to achieve safe operation between internet and the district education LAN (intranet), network firewall and invasion detection are required as basic safety control, however, we also clearly know that the biggest security hidden danger often comes from insiders, for insiders know more about the operating condition, procedure and bugs and holes of the system than outsiders, that's why we must classify and assign role strategy according to the credibility of insiders, to make sure only minimum operation privileged could be obtained, thus to achieve access control. However, in case of breaking through basic defensive guarantee, how to evaluation loss, repair Loopholes and obtain evidence of harmful intruders as soon as possible are also of the top concerns of developing the System. The framework of audit system please refer to Fig.1.

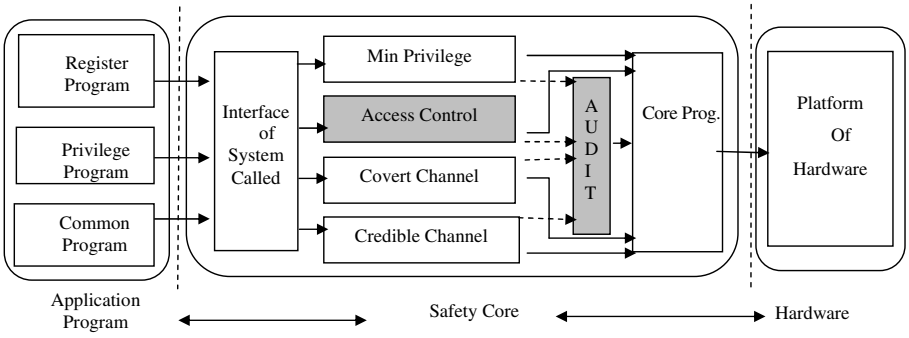


Fig. 1. The audit framework of teaching & researching management

3 Design and Realization of Audit Model in System

3.1 Design of RBAC Basic Security Access Control

The term RBAC is the abbreviation of Role Basic Access Control, which has been widely applied in multiple information systems [4] [5] as one of the research emphasis in access control. According to the minimum access permission control concept, the primary working principle of this model is to classify Users to specified Roles with corresponding Privilege, and Users could activate their Roles by starting some sort of Session, that's the U-R-S-P model in the System.

3.2 Proposal of the Audit Model under RBAC

Basic security of access control could be achieved in developing the System according to a RBAC model, yet it cannot deal with exceptions such as breaking through basic security control, so we propose the audit control strategy under RBAC to enhance security of the system.

As the last security barrier of information system, audit is one of the post-trace analyses, for this reason, audit and analysis of users' pre-operation relies on user's behavior trace kept by system. Such as user behavior log files.

The proposed audit under RBAC model is to modify common U-R-S-P model into the new U-R-S-A-P model. The A means a new independent auditing module was introduced into the RBAC model, meanwhile, the system security supervisor has no right to use other system function, yet behavior of all the other users of the information system would be effectively checked.

3.3 Designing of Auditing Module under RBAC Model

Because the security level of the network teaching & researching management system is classified as level 1 namely Independent protection level, in order to reduce the workload of system security supervisor, the audit log file structure and relevant audit events were simplified for the T & R management system supervisor. But the structure of audit log file could be described as event ID, user ID, occurring time, and events description, event from IP, user's role, original value, and final state.

In out system there are 6 types auditing events were consisted. Those are user, data, transmission, output, peripherals, and remote access. The audit objective is realized by creating 3 public classes of files such as Events, Logs and Response Analysis. In case of an event was activated by user, the first step is calling the function of *Discriminant()* to recognize the event by the Event class, and to see if audit is required. If it need audit, then update the log files by calling corresponding *Insert()* function in the Logs class, and automatically updates the log file database, after that, analyze the event by automatically calling the *Crisis_Resp()* function in the Response class and generate responding alert information to supervisor, and finally modify other correlative information records.

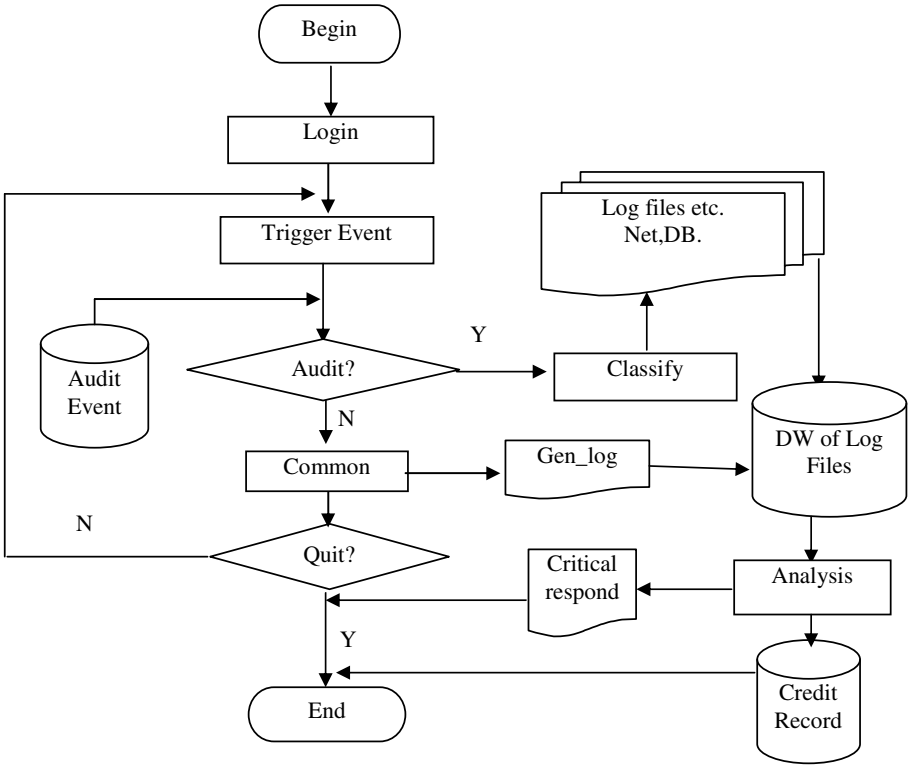


Fig. 2. The flow of audit on RBAC

3.4 Audit Data Flow Analysis under RBAC Model

The comprehensive audit on the System requires 4 log files, i.e. network operation log, database access log, general log and audit log, a system audit database would be obtained by uniformly Extracting Transforming and Loading (ETL operation) the data format of the required contents. We could provide warning strategy, which is based on crisis alarm level, to the system security supervisor by different audit patterns (manually or automatically) conducted on the audit data warehouse and update the

potential credibility of inside operators through typical *Apriori* algorithm of associative rules, for specific RBAC audit model system procedure please refer to Fig.2.

4 Summary

The security system based on RBAC audit model proposed in this article, effectively avoids the possible leaks due to the covet channels meanwhile, the complete log system ensures that the system security supervisor could promptly and precisely locate the leaks and recover the system data through logs operation, the corresponding analysis of the deeper log files in database provides conference for determining insider's credibility, thus to effectively enhance the reliability, perceptiveness and completeness of the Teaching & Researching System.

References

1. Liu, H.: Study on Several Key Technologies in Safety Operation System. Graduate School of Chinese Academy of Science, 1–5 (2002)
2. State Bureau of Quality and Technical Supervision of China: Information Security Technology Operating System Security Assessment Criteria, pp. GB20008–2005 (2005)
3. China Internet Network Information Center, `http://www.cnnic.net.cn`
4. He, S.: Misusage Detection Model Based on RBAC Database. Xi'an Industrial University Journal 4, 365–370 (2009)
5. Fu, P., Duan, P.: An auditable Role Access Control Model. Chongqing Institute of Industry Journal 1, 1–5 (2008)
6. Sun, B., Hao, S., Wang, Z.: Design and Realization of the Privilege Model in Teaching & Research System. Journal of Hebei North University 4, 57–62 (2010)

Strengthening the Ability of Applied Writing for Undergraduate Major in Computer in Practice Teaching System

Yinglin Qin[1,*] Fenglian Zen[2], Baoshan You[1], and Zhong Wang[1]

School of Computer Science&Technology, Shandong Jianzhu University, Jinan, China
Department of Shuangcheng, Jinan Nanjiao Hotel, Jinan, China
sdulaoq@126.com, {qzdd2005,youbaoshan,doctorlyly}@163.com

Abstract. Undergraduates lacking of applied writing ability is a common phenomenon. Reasons is analyzed in this thesis. A reform to improve the students' applied writing ability during the practice teaching is proposed. The applied writing ability is trained gradually through five aspects throughout the practice teaching system. They are course experiments, curriculum project, specialty training, graduation project report and the training of teachers. We take the students major in computer science and technology as an example to carry out the reform in recent two years. We find that the students' applied writing ability is improved greatly and it can be realized through a bit change in practice teaching.

Keywords: applied writing; practice teaching system; undergraduate.

1 Introduction

In recent years, as the development of the society, applied writing has been used more and more widely and frequently. Applied writing has the functions of exchanging the information, communicating and coordinating, ordering the management, public education and certification etc. The application of the applied writing refers to all the aspects of social life and various professions. Applied writing is the basic ability of the undergraduates[1-6]. Meanwhile, it is also a basic ability to confront the employment competition and adapt to the development of the society. As a college student majored in computer science & technology, besides the high level of technology, he needs to have strong capability of language expression, which is reflected by applied writing ability. Nowadays, high education have paid attention to the students' specialized technology ability, but applied writing ability hasn't been taken seriously by most of University of Technology. From our own teaching experiences and communication with the enterprises, we feel that there is contradiction between undergraduates' lacking of the ability and enterprises' high request. An undergraduate majored in computer science and technology has many opportunities to do applied writings, such

* This research is supported by Projects of Education Department in Shandong Province, China (No.J08LJ26).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 189–194, 2011.

as: he needs to write the test report, the curriculum project instruction, the software instruction, etc. However, under current educational system, we tend to neglect the important part. Therefore, if applied writing is put into the whole practice teaching system, and is given special attention in teaching process, the undergraduates' writing ability would be improved. At the same time, the undergraduates learn applied writing through examples. This also avoids the empty talk of applied writing. So this method has more advantageous in training the undergraduates' applied writing ability.

The widespread truth of undergraduates' lacking of applied writing ability is analyzed in this thesis. The reason why it is caused is also analyzed. The solutions of the problem are pointed out. Undergraduates' applied writing ability is contrasted between pre-reform and reform to illustrate the effectiveness of training applied writing during practice teaching.

2 Current Problems

It has been pointed out that applied writing ability is the new request for the university education. So it needs to be taken seriously in the university of technology. The factors contributed to the undergraduates' lack of applied writing ability are from the students, from the professors or lecturers and form the university. We take undergraduates major in computer and technology as an example to analyzed the problems.

2.1 Students Attaching No Importance to Applied Writing

Currently, some universities of technology emphasize engineering courses and neglect the importance of applied writing. On the one hand, students' applied writing is very week, and the writings are not normative. On the other hand, because of professors' or lecturers' conducting problems, students dislike learning the method of applied writing. For example, students in our university are asked to operate following the instruction and write down the data in the experiments. But they have no idea on how to write down the data effectively, how to manage the data and what the results should get from the data. They rely on the teachers' reference book. Their writings are mainly copies of the reference books without their own idea. If there isn't referring data figure in the referencing book, their writing usually tend to be disordered and inexpressive. As the influence of the whole education system that emphasize technology, students tend to think that there is no difference whether to learn applied writing well or not. They also think that technology is more useful. So they have a perfunctory attitude towards test reports, practical reports and even the graduation project report.

2.2 Teachers Not Taking Applied Writing Seriously

Teachers act as guiders in teaching activities. Although teachers work very hard in passing on their technical stills and got good effect, as professionals, most teachers haven't taken applied writing seriously. From the spot-checking of our university in recent years, the problems that due to the teachers are mainly lied in two aspects. For one thing, teachers don't come up with the solutions that can solve students' reports problem and the copies. Teachers always told the students the expression of test results in advance. They even ask the students to express the results according to their request.

Thus, this can exercise some students' standard expressing ability. But it is very import whether the teachers' expression is suitable. For another thing, the teachers don't give students enough constructions in applied writing. This is partly because, the teachers themselves don't have well educated in this field. There are still a few teachers have perfunctory attitudes towards this. They think that the specialized practical teaching means to teach the students the technology only. Teachers' attitude can surely affect students.

2.3 Universities Neglecting Applied Writing in the Overall Training System

Now we can see from the domestic universities especially from the universities of technology, few ones set applied writing courses as a basic course, and few ones set it as a public elective course. In the universities that set applied writing course up, the course and the teacher who teaches this course haven't been emphasized. They thought there is no difference whether set up this course. In the few universities that have set up this course, there is still little effect. One reason is students are unwilling to study. The other reason is that the content of the course is too wide. It hasn't set up different content according to the different circumstances, so it can't meet different majors. In addition, the teachers who teach this course are usually the ones whose major in liberal arts. They have no professional background, or they are not familiar with engineering technology writing. So the teaching contents are impractical. Currently, in our university educational system, applied writing hasn't been a basic course or a public elective course. Undergraduates majored in computer sciences need to improve their applied writing.

To analysis undergraduates' the ill ability of applied writing comprehensively, we can find out that there students' factor, teachers' factor as well as the university's factor. Under current circumstances, it is urgent to find a method to improve students' applied writing ability in the educational system.

3 Strategy : Applied Writing Ability Training in Practical Teaching

In order to solve the problem of students' lack of applied writing ability, we believe it is difficult to add the course to the educational system directly. From the colleges that have the course, the effect is not favorable. It is mainly because as a course, applied writing feels boring. Students don't interest in it. We think that students majored in computer science and technology have many opportunities to involve in applied writing, such as experiments for each course, the curriculum project, specialty training, graduating project, etc. If students are taught and led how important the applied writing are, students' writing ability can improve greatly. Meanwhile, we can strengthen the management of the teachers and lead them to discover and carry out the applied writing through the practice teaching system. So we may divide applied writing ability training into five factors for the undergraduates majored in computer science.

3.1 Executing Applied Writing Ability Training through Course Experiments

There are experimental courses in each basic courses and specialized courses in the process of educational system of students majored in computer. These experimental courses include software courses and hardware courses. In addition to that, there are courses combined in hardware and software. If the experiment report writing can be led during the courses, students can have a new insight into the three kinds of reports and they will be well prepared in the future. What's more important, through the exercises, the students can learn the importance and usefulness of applied writing, further more they may be interested in it. The teachers should be stricter with their students through the process. There may be some difficulties at the beginning, but if the process can keep on a long time, there must be effective.

3.2 Executing Applied Writing Ability Training through Curriculum Project

In the practical teaching system, there are usually 4 or 5 courses of curriculum practical teaching segments to check out students' comprehensive knowledge of each course. The establishment of curriculum project can not only check out the students' knowledge and experiment skills but also the applied writing skills. To train the ability of applied writing through curriculum project lays the foundation for the students to write reports in the future. Undergraduates can choose the subject that offered by the teachers or choose the subject themselves. Through writing, students can use comprehensive knowledge and write down the correct documents. Therefore, the curriculum project writing is very important in applied writing. Students need to carry on the generalized analysis of the algorithm design and projects and get the conclusion. When writing curriculum project, students need to think of the proper expression for each part. For instance, when design the project, students should know how to express clearly and effectively. Through this comprehensive process, students' applied writing ability can be improved.

3.3 Executing Applied Writing Ability Training through Specialty Training Teaching

Specialty training teaching is a comprehensive teaching link in students majored in computer science and technology before they make the graduation project. It trains the students by a model which is similar to a real program. In the training course, we let the practical teaching tutors to train the students. Apart from the practical teaching task, these tutors have no other teaching assignments so as to make the specialized practical training have a good effect. For example, during the embedded system practical training, students are required to analyses the topic and write down the report, then they are required to write the suitable analyzing report, which includes hardware and software design. Further more, they are required to refresh the project design during the process of design. At last, they are required to write down the debugging report, which is mainly the conclusion of the problems that met in the process. When the debugging process is over, students are required to draw up the instruction. We can combine reports of each stage to make them as a specialized practical report. Through this process, students have a clear impression of the applied writing in the project. Their applied writing ability has been improved.

3.4 Executing Applied Writing Ability Training through Graduation Project Design

The graduation project design is the final link of the educational system as well as an important segment of students' comprehensive ability training. By the graduation project, students can adapt to the job easily. So, it is needed to exercise students wholly in this segment. The topic of the graduation project is usually a large one that requires the several persons to do together. Through specialized practical training, students' specialty knowledge has improved a lot, but they have little exercise on enterprises' assignments. We hope the training can be carried on in this segment. Under the cooperative mode, each one has his own task. Each task is relatively independent, but there are mutual connections between them. The mutual connections are the instructive words that bond the students. Through these instructive words, we can find out that which task has no person. Meanwhile, we can also learn other people's task from the instructive words. So the writing material is not for a certain person but for all the members in the project, and thus the instructive words should be special. Through the lead in the graduation project and exercises from former stages, it is easy for the students to master it. At the same time, the writing of graduation project instruction is an important link. So it is necessary to be strict with students' writing style. Only in this way, can the students master the ability.

3.5 Attaching Importance to Teachers' Further Education and Strengthening Teachers' Management

In the former four stages, teachers are a decisive factor in the process of training the applied writing. The teaching teacher is not the specialized writing teacher. Their applied writing level determines the training of applied writing. So the teachers' applied writing ability should be strengthened by continuous teaching activities and collective preparations. Meanwhile, teachers' responsibility managements should be enhanced. Resign some teachers who are not strict with students. It is possible to realize the practical training by practical teaching.

Since 2009, we have made a plan for undergraduates majored in computer science and technology to improve their applied writing ability. In our daily teaching activities, we pay attention to train their writing skills. The effect is favorable. From the top-check of experiment reports, curriculum project reports and graduation project reports, students' ability have improved greatly. Most of the graduates think applied writing is helpful for their future job.

4 Conclusion

Applied writing ability is an important component of students' comprehensive abilities. Reasons are analyzed for students' lack of applied writing ability in university of technology. It also put up with some methods to improve students' applied writing ability. It is divided into 5 parts to instruct the methods. They are course experiments, course design, specialty training, graduation project report and the training of teachers. We take the students major in computer science and technology as an example to carry out the reform in recent two years. We find that the students' applied writing ability is improved greatly and it can be realized through a bit change in practice teaching.

References

1. Zhang, X.: Application of the Situational Approach to Teaching the Course Applied Writing. Journal of Xi'an Eurasia University 8(3), 80–82 (2010) (in Chinese)
2. Sun, L.: Problems Existing in the Teaching of Applied Writing and Countermeasures. Journal of Tianjing Vocational Institutes 1(12), 87–89 (2010) (in Chinese)
3. Wu, R.: Probe into open teaching of applied writing in independent. Theory and Practice of Education 28, 56–58 (2008) (in Chinese)
4. Xuan, A.: Exploration and Practice of the Teaching System of Applied Writing. Journal of Nanyang Normal University(Social Sciences) 7(11), 78–79 (2008) (in Chinese)
5. He, J., Xu, L.: Application of Situation teaching to Writing Class. Theory and Practice of Education 25, 47–48 (2005) (in Chinese)
6. Zeng, R.: Employment-oriented Training of Writing Skills. Higher Education Forum 2, 66–68 (2009) (in Chinese)

An Error Location and Diagnosis Method for Communication Test of High-Speed Railway Train Control System Based on String Pattern Matching

Liu Huang, Wei Dong*, Yindong Ji*, and Xinya Sun

Tsinghua National Laboratory for Information Science and Technology
Department of Automation, Tsinghua University
Beijing 100084, China
huanguil@gmail.com, {weidong,jyd,xinyasun}@tsinghua.edu.cn

Abstract. Train-ground communication data that are generated during the operation of high-speed railway train control system are key information which reflects the safety control logic and function of high-speed railways. Since the possible existence of defects in the devices of train control system, the communication data may contain some errors (e.g. missing, redundant or wrong messages). As a result, error location and diagnosis of these communication data is an important part of function test in train-ground communication. These problems can be abstracted into an approximate string matching problem which has seldom been studied in previous research. This paper extends the bit-parallel algorithm for approximate regular expression matching problem and proposes an online method to locate and diagnose errors. Finally, a case study is presented and the results indicate the effectiveness of this method.

Keywords: train-ground communication test; error location and diagnosis; approximate regular expression matching; bit-parallel algorithm.

1 Introduction

Chinese Train Control System Level 3(or CTCS-3) is designed for Chinese high-speed railways, which aims to provide a speed of over 300 km/h and meanwhile ensure the safety. The normal and reliable operation of CTCS-3 mainly depends on the two-way continuous communication between train and ground. Therefore, the test for train-ground communication function of newly developed train control system is a vital necessity. The communication data produced in the test are a good reflection of the communication logic and function, thus error location and diagnosis of these data is an important part of the communication function test.

The communication data are composed of "messages" and their correct exchanging order is defined by the communication specification. So an error can be defined as missing, redundancy or replacement of a message. Under this definition, fault monitoring task includes locating the error's position in the actual communication data and diagnosing the type of the located error.

* IEEE Member.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 195–204, 2011.

If a message is seen as a “character”, then the data from actual operation can be seen as a “text” in string format and the communication specification a “pattern”. The “text” contains certain substrings similar to the “pattern”, but with a few differences which need locating and diagnosing. From this point of view, the problem of locating and diagnosing errors in train-ground data can be abstracted into an approximate string matching problem.

However, the existing algorithms mainly focus on searching substrings close to the pattern in the text. 0 proposes a dynamic programming algorithm for approximate string matching, which is the basis of later algorithms. A method based on deterministic finite automaton is presented in 2, which imitates the searching process in text. Bit-parallel method is efficient in solving approximate matching problem and it can be applied in the form of row-wise bit-parallelism 3 and diagonal-wise bit-parallelism 4. 5 describes a bit-parallel algorithm which uses a series of differences of dynamic programming matrix. Approximate regular expression matching extends the pattern from a single string to a regular expression and it is the most complicated circumstance in approximate matching. 3 and 6 pose extension to previous bit-parallel algorithm so that their algorithms can be adapted to approximate regular expression matching.

To sum up, the currently available algorithms for approximate matching only determine whether the text contains substrings close to the pattern. But compared with the pattern, the substring has some errors such as the insertion, substitution or deletion of some characters which need locating and diagnosing. Especially when the pattern is a regular expression, it is necessary to find an algorithm to locate and diagnose these errors because errors are not easy to discover under such circumstance. However, little research has been carried out in this aspect. This paper extends the bit-parallel algorithm of approximate regular expression matching and presents an online method to locate and diagnose errors. This method can be applied to cases with a single error or multiple separate errors.

This paper is organized as follows. In Section 2, the model of string matching problem is introduced. An error locating and diagnosing method based on approximate regular expression matching is presented in Section 3. Section 4 proposes an application case to indicate this method can locate and diagnose the errors in train-ground communication data effectively. Section 5 concludes the paper and provides some suggestions for future research.

2 Basic Model of String Pattern Matching

2.1 String Matching and Regular Expression Matching

A string is a character sequence defined on an alphabet Σ. Single string matching is to search a pattern $p = p_1 p_2 \ldots p_m$ as a substring in a text $T = t_1 t_2 \ldots t_n$.

Regular expression matching is a kind of string matching with a regular expression as the pattern. Given a regular expression *RE*, the language represented by *RE*, or *L(RE)*, is a set of strings on Σ. The prefix set of *RE* is defined as $pre(RE) = \{v \mid \exists w \in L(RE), v \in pre(w)\}$, and $pre(w)$ is the prefix set of string w.

2.2 Approximate String Matching

Approximate string matching, or "string matching allowing errors", is to find in text T a substring which permits k differences (it means k differences or fewer) from the pattern p.

We adopt a commonly used model called *edit distance* 7 to measure a difference. This model defines a difference as an edit operation of one character including insertion, deletion, or substitution. Moreover, the edit distance *ed(x,y)* between strings x and y means the fewest edit operations needed to make x equal to y. By using the edit distance model, the approximate string matching can be described as a problem of searching every p' which satisfies $ed(p, p') \le k$ in text T 8.

2.3 Approximate Regular Expression Matching Using Bit-Parallel Algorithm

Approximate regular expression matching is an extension of approximate string matching. Given the regular expression *RE*, the problem of approximate matching allowing k errors can be stated as follows: find in text T every substring w' satisfying the condition that $\exists w \in L(RE), s.t. ed(w, w') \le k$.

The so called bit-parallel algorithm handling this problem can be presented below 08.

First of all, the regular expression *RE* should be transformed into a nondeterministic finite automaton (NFA). Here, we choose a most popular way that is called *Glushkov construction* 9 to build the NFA.

The corresponding NFA N can be represented as a quintuple $N = (Q = \{s_0 \ldots s_m\}, \Sigma, I = s_0, F, \Delta)$ with m+1 states if *RE* has m characters belonging to Σ.

Bit-parallel algorithm views an array of boolean values as bits of a computer word (these bits can also be called "*bit mask*") so that these values can be updated by a single bitwise operation of the computer word. If the active state set of NFA is expressed in a bit mask, this bit mask can renew itself via a bitwise operation instead of traversing through the whole set.

A bit mask R_0 of length m+1 is used to denote the active state set as NFA N has m+1 states. R_0^j denotes the active state set after j-th character in the text is processed. And more precisely, $R_0^j[i] = 1$ means that s_i belongs to the active state set while $R_0^j[i] = 0$ means the opposite. The initial value of R_0 is $R_0^0 = 0^m 1$. The final state set represented in bit mask is $F_n = \left|_{s_j \in F} 0^{m-j} 10^j\right.$. The transition set Δ can be expressed by table B_n, where

$$B_n[i, \sigma] = \left|_{(s_i, \sigma, s_j) \in \Delta} 0^{m-j} 10^j\right. \tag{1}$$

represents the bit mask of states within reach from state i by character σ.

A specific property of the NFA deriving from Glushkov construction is that all the arrows leading to a given state are labeled by the same character 10. For this reason, table $B[\sigma]$ and $T_d[D]$ can be used to calculate transitions, where

$$B[\sigma] = \Big|_{i \in 0 \ldots m} B_n[i, \sigma] \tag{2}$$

indicates the bit mask of states reachable by character σ from any state, and

$$T_d[D] = \Big|_{\{(i,\sigma) \mid D \& 0^{m-i}10^i \neq 0^{m+1}, \sigma \in \Sigma\}} B_n[i, \sigma] \tag{3}$$

indicates bit mask of states reachable from D(D is a given bit mask of length m+1) by any character in Σ.

The formula of updating R_0^j is:

$$R_0^{j+1} = \left(T_d[R_0^j] \,\&\, B[t_{j+1}]\right) \mid 0^m 1 \tag{4}$$

where the bitwise OR represents the self loop of the initial state in the NFA which can be passed by any character. In this way, the language of NFA is converted into $\Sigma^* L(RE)$ from $L(RE)$. A successful matching (exact matching) is reported when the condition $R_0^j \,\&\, F_n \neq 0^{m+1}$ is met.

Suppose k errors are allowed in approximate matching. Similarly, bit masks $R_d^j (d = 1 \ldots k)$ of length m+1 are used to denote the active state set allowing d errors up to t_j in text. The equation $R_d^j[i] = 1$ is true if and only if s_i belongs to the active state set up to t_j with $\leq d$ errors. The mathematical explanation is that $\exists v \in pre(RE), \exists h < j$ satisfying the condition that $ed(v, t_h \ldots t_j) \leq d$ and the end state is s_i after v is executed in N from s_0.

Before searching the text, these bit masks are initialized with $R_0^0 = 0^m 1, R_d^0 = T_d[R_{d-1}^0] \mid 0^m 1 (d = 1 \ldots k)$. The updating formulas are:

$$R_0^{j+1} = \left(T_d[R_0^j] \,\&\, B[t_{j+1}]\right) \mid 0^m 1 \tag{5}$$

$$R_d^{j+1} = (T_d[R_d^j] \,\&\, B[t_{j+1}]) \mid R_{d-1}^j \mid T_d[R_{d-1}^j] \mid T_d[R_{d-1}^{j+1}] \quad (d = 1 \ldots k) \tag{6}$$

where R_d^{j+1} has four terms representing matching, inserting, substituting and deleting a character in the pattern respectively. A further explanation of these four terms is stated in 0. A successful approximate matching allowing d errors may happen when the condition $R_d^j \,\&\, F_n \neq 0^{m+1} (d = 1 \ldots k)$ satisfies.

3 Algorithm of Error Location and Diagnosis

Given the regular expression *RE*, previous approximate matching algorithms only provide all substrings *w'* in text *T* which meet the condition that $\exists w \in L(RE), st.ed(w, w') \leq k$. On this basis, our algorithm can further locate the errors occur in *w'* and diagnose the type of errors (insertion, substitution or deletion) online.

Firstly, let us discuss the simplest situation. Suppose k=1, which means that compared to the pattern w, the substring w' can only allow one insertion, substitution or deletion.

3.1 Error Location

In the process of searching text T, if no matching occurs between the text (currently read in part, similarly hereinafter) and RE, i.e. any suffix of the text does not match any $v \in pre(RE)$, then R_0 remains 0^m1.

When there exists a suffix v' of T matching $v \in pre(RE)$, some of the first m bits of R_0 are set to 1, which indicate the active states in the NFA currently.

If an error appears between v' and v with more characters being read in, R_0 returns to 0^m1. According to the above analysis, we can have the method to locate the error.

Theorem 1. When $R_0^{j-1} \neq 0^m1$ and $R_0^j = 0^m1$, there must exists an error at t_j.

3.2 Error Diagnosis

Insertion. Insertion is to insert a character in the pattern w (w is a string and $w \in L(RE)$) so that substring w' is the same as w.

Theorem 2. The sufficient and necessary condition of t_j being an insertion is that an error is located at t_j and that the inequality $T_d[R_0^{j-1}] \& B[t_{j+1}] \neq 0$ is satisfied.

Proof. Sufficiency: As $T_d[R_0^{j-1}] \& B[t_{j+1}] \neq 0$ implies, the states reachable through one transition from the state set R_0^{j-1} have intersection with the states that can be reached by character t_{j+1}. And one property of Glushkov automaton is that all the arrows arriving at a given state are marked with the same character, therefore the states reachable by character t_{j+1} cannot be reached by any other character. That is to say, t_{j-1} should be followed by t_{j+1}. In addition, an error is detected at t_j, thus t_j is an inserted character.

Necessity: As is known that t_j is an inserted character, so the character after t_{j-1} should be t_{j+1} rather than t_j. From this, we have $T_d[R_0^{j-1}] \& B[t_{j+1}] \neq 0$ and $T_d[R_0^{j-1}] \& B[t_j] = 0$. And $R_0^j = \left(T_d[R_0^{j-1}] \& B[t_j]\right) | 0^m1 = 0^m1$ indicates that an error is located at t_j. □

Substitution. Substitution is to change a character in the pattern w (w is a string and $w \in L(RE)$) into a character in substring w' so that w' is the same as w.

Theorem 3. The sufficient and necessary condition of t_j being an substitution is that an error is located at t_j and that the inequality $T_d[T_d[R_0^{j-1}]] \& B[t_{j+1}] \neq 0$ is satisfied.

Proof. Sufficiency: As $T_d[T_d[R_0^{j-1}]]\&B[t_{j+1}]\neq 0$ implies, the states reachable through two transitions from the state set R_0^{j-1} have intersection with the states that can be reached by character t_{j+1}. That is, the second character after t_{j-1} should be t_{j+1}. In addition, an error is detected at t_j, thus t_j is an substituted character.

Necessity: As is known that t_j is an substituted character, so the character after t_{j-1} should not be t_j, and that the second one after t_{j-1} should be t_{j+1}. From this, we have $T_d[T_d[R_0^{j-1}]]\&B[t_{j+1}]\neq 0$ and $T_d[R_0^{j-1}]\&B[t_j]=0$. And $R_0^j=\left(T_d[R_0^{j-1}]\&B[t_j]\right)|0^m1=0^m1$ indicates that an error is located at t_j. □

Deletion. Deletion is to delete a character in the pattern *w* (*w* is a string and $w\in L(RE)$) so that substring *w'* is the same as *w*.

Theorem 4. The sufficient and necessary condition of a missing character between t_{j-1} and t_j is that an error is located at t_j and that the inequalities $T_d[T_d[R_0^{j-1}]]\&B[t_j]\neq 0$, $T_d\left[T_d[R_1^{j-1}]\&B[t_j]\right]\&B[t_{j+1}]\neq 0$ are satisfied.

Proof. Sufficiency: As $T_d[T_d[R_0^{j-1}]]\&B[t_j]\neq 0$ implies, the second character after t_{j-1} should be t_j. And $T_d\left[T_d[R_1^{j-1}]\&B[t_j]\right]\&B[t_{j+1}]\neq 0$ means that the character after t_j should be t_{j+1}. In addition, an error is detected at t_j, which means that t_j cannot be the character following t_{j-1}. Thus, there is a missing character between t_{j-1} and t_j.

Necessity: As is known that there is a missing character between t_{j-1} and t_j, so the second character after t_{j-1} should be t_j. From this, we have $T_d[T_d[R_0^{j-1}]]\&B[t_j]\neq 0$ and $T_d[R_0^{j-1}]\&B[t_j]=0$. And $R_0^j=\left(T_d[R_0^{j-1}]\&B[t_j]\right)|0^m1=0^m1$ indicates that an error is located at t_j.

Furthermore, in case of mistaking insertion or substitution for deletion, the character after t_j should be ensured as t_{j+1}. That is, the condition $T_d\left[T_d[R_1^{j-1}]\&B[t_j]\right]\&B[t_{j+1}]\neq 0$ should be satisfied. □

3.3 An Extension to the Multiple Error Case

The above four theorems can all be extended to a multiple error case. Here we only discuss the case with multiple separate errors. Two or more successive errors are not considered in this paper. The proofs are omitted as they are similar to Theorem 1-4.

Theorem 5. When $T_d[R_d^{j-1}]\&B[t_j]\neq 0$ and $T_d[R_d^j]\&B[t_{j+1}]=0$, there must exists an error at t_j, which is the *(d+1)*-th $(d=0,\ldots k-1)$ error in the latest matching process.

Theorem 6. The sufficient and necessary condition of t_j being an insertion is that an (d+1)-th $(d = 0,\ldots k-1)$ error is located at t_j and that the inequality $T_d[R_d^{j-1}] \& B[t_{j+1}] \neq 0$ is satisfied.

Theorem 7. The sufficient and necessary condition of t_j being an substitution is that an (d+1)-th $(d = 0,\ldots k-1)$ error is located at t_j and that the inequality $T_d[T_d[R_d^{j-1}]] \& B[t_{j+1}] \neq 0$ is satisfied.

Theorem 8. The sufficient and necessary condition of a missing character between t_{j-1} and t_j is that an (d+1)-th $(d = 0,\ldots k-1)$ error is located at t_j and that the inequalities $T_d[T_d[R_d^{j-1}]] \& B[t_j] \neq 0$, $T_d\left[T_d[R_{d+1}^{j-1}] \& B[t_j]\right] \& B[t_{j+1}] \neq 0$ are satisfied.

4 Application

4.1 Problem Description

CTCS-3 can be divided into two subsystems: ground subsystem and onboard subsystem. The ground subsystem includes Radio Block Center(RBC), balise, track circuit, GSM for railways(GSM-R), etc, while the onboard subsystem includes automatic train protection(ATP), onboard GSM-R module, etc. CTCS-3 is based on two-way wireless communication between train (mainly ATP) and ground (mainly RBC) through GSM-R. ATP and RBC exchange data in the form of "messages".

This problem can be solved using the method in Section 3 via an abstraction. Since a missing, redundant or wrong message corresponds to deletion, insertion or substitution respectively in approximate string matching, all we need to do first is transforming the specification into a regular expression or an automaton.

4.2 Specification for Train-Ground Communication

The specification for train-ground communication which sets the standard order of messages is described in UML sequence diagrams. Fig. 1 gives a possible communication standard between train and ground in the scenario "mission start" 11. (In Fig.1, M represents message and P represents package which is a part of a message. The number next to M or P indicates the ID of message or package.)

Here, we convert the specification into an automaton. RBC in the form of automaton transformed from Fig. 1 is illustrated in Fig. 2. It is worth noticing that the automaton in Fig. 2 is in accordance with the property of Glushkov's NFA, i.e. all the arrows leading to a given state are labeled by the same character. As a result, the algorithm introduced in Section 3 can be adopted.

4.3 Case Study

This section illustrates some examples of searching the automaton of Fig.2 in a train-ground communication message sequence. The first three examples contain one error while the fourth contains two errors.

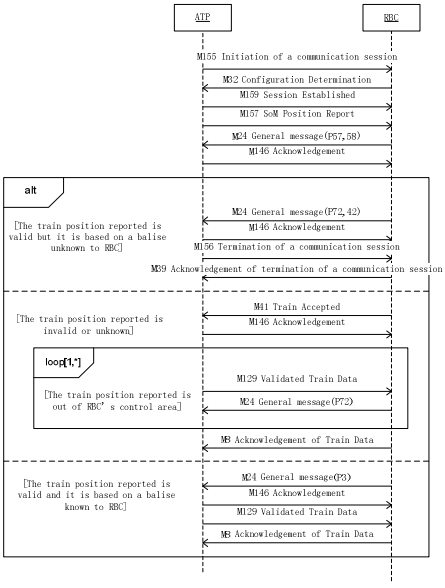


Fig. 1. A possible communication standard between train and ground in the scenario "mission start" in UML sequence diagram

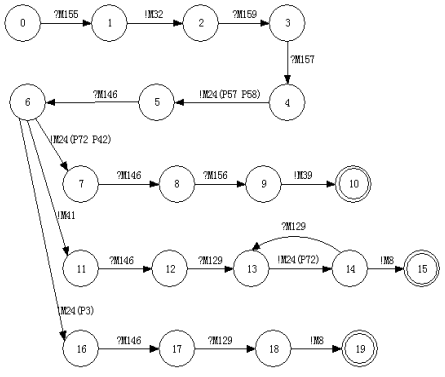


Fig. 2. The automaton of RBC in scenario "mission start"

(1) For a message sequence "155, 32, 159, 157, 24, 146, 146, 41, 146, 129, 24, 129, 24, 8" (the numbers represent the message ID, similarly hereinafter), an error is located after the second "146" is read in, and an "insertion" is reported after "41" is read in, i.e. the message 146 before message 41 is an redundant message.

(2) For a message sequence "155, 32, 159, 157, 146, 41, 146, 129, 24, 129, 24, 8", an error is located after the first "146" is read in, and an "deletion" is reported after "41" is read in, i.e. before the first message 146 there exists a missing message 24.

(3) For a message sequence “155, 32, 159, 157, 24, 146, 41, 146, 129, 129, 129, 24, 8”, an error is located after the second “129” is read in, and an “substitution” is reported after the last “129” is read in, i.e. the second message 129 is a wrong message while the right one is 24.

(4) For a message sequence “155, 32, 159, 157, 146, 41, 146, 129, 129, 129, 24, 8”, the first error is located after the first “146” is read in, and an “deletion” is reported after “41” is read in, i.e. before the first message 146 there exists a missing message 24. And then the second error is located after the second “129” is read in, and an “substitution” is reported after the last “129” is read in, i.e. the second message 129 is a wrong message while the right one is 24.

5 Conclusion

This paper extends the bit-parallel algorithm of approximate string matching for regular expressions and proposes an online method to locate and diagnose errors. Some examples above indicate that this method can be applied to the error location and diagnosis in train-ground communication data.

Our future work can be extended in the following directions:

(1) This paper only discusses the location and diagnosis of several separate errors. And the case with successive errors is a more complicated problem for further study.
(2) Consider a new type of error: transposition of two neighbouring characters which is commonly seen in texts. The method of locating and diagnosing this error type should be investigated.

Acknowledgments. This work was supported in part by the National Key Technology R&D Program under Grant 2009BAG12A08, the Research Foundation of the Ministry of Railways and Tsinghua University (RFMOR&THU) under Grant J2009Z28 and the Research Foundation of Beijing National Railway Research and Design Institute of Signal and Communication.

References

1. Sellers, P.H.: The theory and computation of evolutionary distances: Pattern recognition. Journal of Algorithms 1, 359–373 (1980)
2. Ukkonen, E.: Finding approximate patterns in strings. Journal of Algorithms 6, 132–137 (1985)
3. Wu, S., Manber, U.: Fast text searching allowing errors. Communications of ACM, 83–91 (1992)
4. Baeza-Yates, R.A., Navarro, G.: Faster approximate string matching. Algorithmica 23, 127–158 (1999)
5. Myers, E.W.: A fast bit-vector algorithm for approximate string matching based on dynamic programming. Journal of the ACM 46, 395–415 (1999)
6. Navarro, G.: Nr-grep: a fast and flexible pattern matching tool. Software Practice and Experience 31, 1265–1312 (2001)

7. Levenshtein, V.I.: Binary codes capable of correcting spurious insertions and deletions of ones. Problems of Information Transmission 1, 8–17 (1965)
8. Navarro, G., Raffinot, M.: Flexble pattern matching in strings, vol. 145, pp. 117–123. Cambridge University Press, New York (2002)
9. Gluskov, V.M.: The abstract theory of automata. Russian Mathematical Surveys 16, 1–53 (1961)
10. Navarro, G., Raffinot, M.: Compact DFA representation for fast regular expression search. In: 5th International Workshop Algorithm Engineering, p. 6 (August 2001)
11. Zhang, S.: CTCS-3 general technical scheme, pp. 6–10. China Railway Press (2008)

A Novel Stepper Speed Control Algorithm Based on FPGA

Baoshan You[1] and Liying Pei[2]

[1] ShanDong Jianzhu University, Jinan 250101, China
[2] Jinan Vocational College, Jinan, 250103, China
youbs@sdjzu.edu.cn, peiliying1225@126.com

Abstract. To solve the problems existed in traditional stepper speed control algorithm, a novel algorithm was introduced. This new control algorithm is dramatic different from traditional system, it is based on space control instead of time control. It was implemented by FPGA with VHDL, simulated with Altera Quartus II software and tested by a real pen-plotter control system. The system based on this novel control algorithm have dramatic advantage, which is all critical parameters can be changed as often as possible without system modification. Base on this novel speed control algorithm, a pen-plotter control system was implemented in a FPGA, which is a correct, flexible, high performance, low cost system. This novel stepper speed control algorithm can be used for a wide range of application in industrial automation.

Keywords: Stepper, Speed Control, FPGA.

1 Introduction

Stepper motor is widely used in various applications, such as printers, disk drives, engraving machine, pen plotters, autonomous cars, robotic arms. It is a "digital" version of the electric motor. The rotor moves in discrete steps as commanded, rather than rotating continuously like a conventional motor[1]. It is highly effective in motion control application for high precision and high performance of torque control. it offers better torque performance over wider speed ranges[2]. It has many advantages, just like the high pointing accuracy, low cost, free error of position accumulation, easy to control, and so on[3-4]. One of the most significant advantages of a stepper motor is its ability to be accurately controlled in an open loop system. Open loop control means no feedback information about position is needed. This type of control eliminates the need for expensive sensing and feedback devices such as optical encoders. Here position is known simply by keeping track of the input step pulses. With applications where load torque is stable and operations are at low speed, an open loop driver is preferred since it needs a simple control algorithm and an encoder is not required [5]. However, it has disadvantages, when stepper works at high speed in an open loop system, step missing and overshooting phenomena maybe happened. So speed control becomes an important issue to make ensure stepper works properly. In this paper, the control of speed of the stepper motor is done by using FPGA (Field Programmable Gate Array) instead of DSP & Microprocessor [6].

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 205–209, 2011.

2 Stepper Speed Control

Due to mechanism of stepper motor, if we want to drive motor slewing in high speed, it must be driven in slow speed, increased slowly to high speed in order to avoid missing step, and decreased slowly to avoid overshooting step. There are three widespread speed control profiles: parabolic, trapezoidal, and S-curve. In this paper, trapezoidal profile is used and introduced. It is alternatively called the linear acceleration and deceleration. It can be implemented very easily in different system. Fig. 1 is the traditional linear acceleration and deceleration profile, one axis is time, another is frequency (Pulses Per Second). According Eq. 1, f_n can be calculated very easily, f_{min} and f_{max} are given value by different application. So a linear acceleration and deceleration table can be created. It can be implemented by DSP & Microprocessor with timer. Frequency can be changed with different initial value of the timer. This method was implemented in many books and papers, details are omitted. It has a weakness that the table should be recalculated when the a (acceleration coefficient) changed in Eq. 1. In real system, a is changed often.

$$f = f_{\min} + at \tag{1}$$

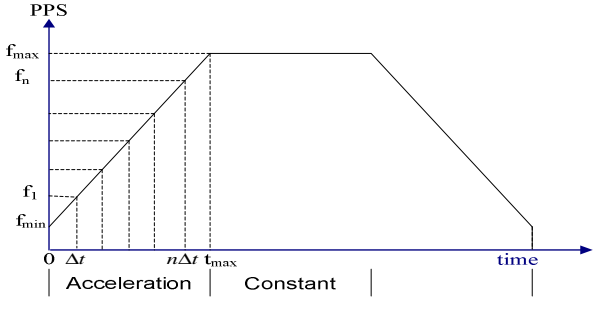


Fig. 1. Traditional linear acceleration and deceleration

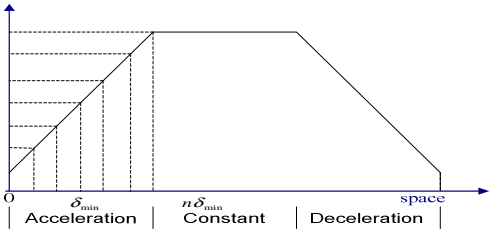


Fig. 2. Novel linear acceleration and deceleration

Fig. 2 is a new method implemented by this paper with FPGA. Fig. 2 and Fig. 1 are very similar, the key difference is horizontal axis, one is time, and the other is space. In

Fig. 2, $\delta_{\min}$ is the minimum pulse equivalent which unit is millimeter per pulse. In this paper, X-Y table system was introduced as Fig. 3. There is δ_x for X axis pulse equivalent, δ_y for Y axis pulse equivalent, $\delta_{\min}$ is for virtual ρ axis. Among of them, $\delta_{\min}$ is the minimum. In this paper, $\delta_{\min}$ =0.0001mm. So pulse number for X, Y and virtual ρ direction to feed can be get by Eqs. 2, the value are P_x, P_y and P_ρ. ρ can be get by $\rho = \sqrt{x^2 + y^2}$. There are there directions to feed pulse, virtual ρ direction need speed control, X and Y directions pulse to feed stepper can be got by linear interpolation to virtual ρ direction. In fact, virtual ρ direction never feed real pulse to stepper. In this paper, a novel stepper speed control algorithm is for virtual ρ direction, the other two direction of X and Y do not need speed control again, because their speed are controlled by interpolation to virtual ρ direction. Eqs. 3, Eqs. 4 and Eqs. 5 are a group of equations to used for speed control, among of them, $v_{\max}$, v_0 and a are specific values given by different system, v is speed, a is acceleration. Eqs.3 are equations for acceleration, Eqs.4 are equations for constant, Eqs. 5 are equations for deceleration. With these three group equations, this paper change time value Δt to space value s. It seems really easy and obviously. In fact, Δt is a key parameter to deal with. There are two factors should be think about, one is accuracy, the other is short enough to follow the stepper movement. In traditional system based on DSP & Microprocessor, Δt can not get accurately, because DSP & Microprocessor work sequentially, their instructions have different execution time, meanwhile, a very short time is a big challenge for traditional system. In this paper, Δt was got by system based on FPGA which can work concurrently. It is really easy to get a Δt, which is short and accurate enough. In this paper, system clock for FPGA system is 50 mega-Hz, 5 clock cycles is a Δt, Δt = 100ns. With the concurrent processing of FPGA, all the calculations based on Eqs.3, Eqs.4 and Eqs.5 can be finished within Δt. If the minimum unit is small enough, which is 10^{-17} in this paper, all the calculations can base on integer with addition and subtraction, multiplication is not necessary again. Due to very small unit, wide registers should be ready to contain the results of calculation.

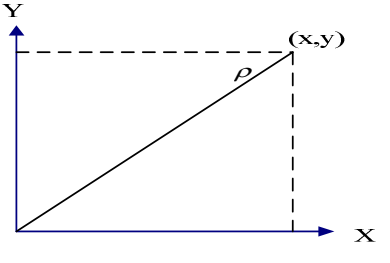


Fig. 3. X-Y table system

$$P_x = \frac{x}{\delta_x} \quad P_y = \frac{y}{\delta_y} \quad P_\rho = \frac{\rho}{\delta_\rho} \tag{2}$$

$$s_i = v_{i-1}\Delta t + \frac{1}{2}a\Delta t^2 \quad v_i = v_{i-1} + a\Delta t \quad s = \sum_{i=1}^{n} s_i \tag{3}$$

$$s_i = v_{i-1}\Delta t \quad v_i = v_{\max} \quad s = \sum_{i=1}^{n} s_i \tag{4}$$

$$s_i = v_{i-1}\Delta t - \frac{1}{2}a\Delta t^2 \quad v_i = v_{i-1} - a\Delta t \quad s = \sum_{i=1}^{n} s_i \tag{5}$$

In Fig. 2, $\delta_{\min}$ and frequency are correlative. This is the focus of this paper. With Eqs. 3, Eqs. 4 and Eqs. 5, s was got with the steppers movement. A register was used to contain s, which was called intSpaceValue in VHDL code for this paper. At the very beginning, s is equal zero, intSpaceValue is equal s .When intSpaceValue is equal or bigger than $\delta_{\min}$, so one pulse should be feed to virtual ρ direction, meanwhile, intSpaceValue = intSpaceValue - $\delta_{\min}$. So this calculation can repeat again. Pulses are feed to virtual ρ direction continuously. At the same time, X and Y directions steppers are feed by pulses got by linear interpolation.

3 Experiments and Results

Altera Quartus II software is used to design and simulated the result of VHDL design. The stepper speed control algorithm is simulated as in Fig. 4. intXSteps and intYSteps are the steps to go in x and y direction, lgcXPulse and lgcYPulse are the pulse output to drive the steppers of x and y direction. The frequency of lgcXPulse and lgcYPulse are changed at different part, the frequency is increasing at the begin part, decreasing in the end part. It is the process of acceleration and deceleration. Except simulation result, a real experiment system based on Altera FPGA EPEC12Q240C8[7] is proved successfully in the real Pen plotter system.

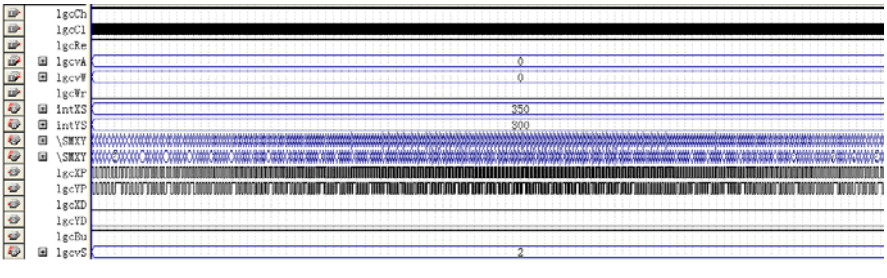


Fig. 4. The simulation result of stepper speed control algorithm

4 Conclusion

In this paper, a novel stepper speed control algorithm based on FPGA was introduced. This control algorithm is dramatic different from traditional system, it is easy to be implemented by FPGA with VHDL. The system based on this novel control algorithm have dramatic advantage, which is all critical parameters can be changed as often as possible without system modification, such as maximum speed $v_{\max}$, acceleration coefficient a, pulse equivalent δ for different axis etc. All this parameters directly influence the performances of the real system. A real system based on this speed control algorithm was implemented in a FPGA chip, Nios II embedded processor[8] was used to construct SoPC(System on a Programmable Chip) with the custom IP(used to implement speed control algorithm with VHDL). It is a correct, flexible, high performance, low cost system. This novel stepper speed control algorithm can be used for a wide range of application in industrial automation.

Acknowledgements

This work was supported by a grant from Shandong Jianzhu University (Project no. XN080106).

References

1. Stepper Motors Guide, http://www.electojects.com/motors/stepper-motors-1.htm
2. Paula, G.: The Rise of VSR Motor. Mechanical Engineering Magazine
3. Bi, S.-X.: Stepper motor drive control of the Applied Research. Tianjin University Master's Thesis, 40–46 (2003)
4. Li, Y.-D.: AC Motor Digital Control System, pp. 152–172. China Machine Press, Beijing (2002)
5. Chen, T.-C., Su, Y.-C.: High Performance Algorithm Realization on FPGA for Stepper Motor Controller. In: SICE Annual Conference, August 20-22. The University Electro-Communications, Japan (2008)
6. Petrov, J., Strmik, S.: A Microcomputer-Based Speed Controller for Lift Drives. IEEE Transactions on Applications 24(3), 487–498 (1988)
7. Altera Corp. Cyclone Device (DB/OL). Handbook, http://www.altera.com.cn/support/devices/cyclone/dev-cyclone.html (August 2005)
8. Altera Corp. Nios II Processor Reference Handbook (DB/OL) (May 2006), http://www.altera.com.cn/literature/lit-nio2.jsp

The Sensitivity Characteristics of the Electrostatic Coupling Intra-Body Communication Based on the Mach-Zehnder Electro-Optic Modulation

Yong Song*, Kai Zhang, Bangzhi Kang, Ruxin Xie, Fuzhou Shang, and Qun Hao

School of Opto-electronic, Beijing Institute of Technology,
Beijing, 100081 China
{yongsong,20040491,20904106,xieruxin,
2120100657,qhao}@bit.edu.cn

Abstract. The Intra-Body Communication (IBC) based on optoelectronic modulation is believed as an available method to achieve high speed and reliable signal transmission based on human body channel. In this paper, the sensitivity characteristics of the electrostatic coupling IBC based on Mach-Zehnder (M-Z) electro-optic modulation have been investigated. Firstly, the mathematic model of the electrostatic coupling IBC based on M-Z optoelectronic modulator has been proposed, while the sensitivity characteristics of this kind of IBC are discussed. Secondly, the corresponding measurements were carried out to verify the proposed model. Our investigation indicates that the simulation results of the sensitivity characteristic coincide with the corresponding in vivo measurement results, while the proposed model will help to achieve reliable, high speed and low cost intra-body communication.

Keywords: Intra-body communication, electrostatic coupling, optoelectronic modulation, personal area network.

1 Introduction

Intra-body communication (IBC) is a technology using human body as a transmission medium for electrical signals [1]. Compared with the wireless communication, it has several novel characteristics, which include high information security, low energy consumption [4], network access based on contacting, [2], etc. Therefore, it is believed that IBC technology will offer the significant advantages in personal area network (PAN) [1], [2], implant biomedical monitoring [3] and interaction between human and environments [5], etc.

Electro-optic (EO) modulation, which is believed as an available method to achieve the reliable and high speed IBC [2], [5], [6]. However, although the EO

* Received his PhD degree from Beijing Institute of Technology, China in 2004. He has been an associate professor at the School of Information Science and Technology, Beijing Institute of Technology, China since 2007. He was a visiting scientist of the Institute of Optics, University of Rochester, America from Oct. 2009 to Apr. 2010. His research interests include intra-body communication, opto-electronic detection, wireless multimedia sensor network.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 210–216, 2011.

modulation method has been used in the IBC field successfully and the high seed of 10 Mbps was achieved, the relationship between the EO sensor's parameters and the sensitivity characteristics of the IBC based on this method has not been gotten fully investigation. Therefore, it's difficult to determine the corresponding parameters of the EO sensor used in IBC by using theoretical method. In this paper, the mathematical model of the IBC system based on M-Z modulator is derived firstly, while the sensitivity characteristics of the electrostatic coupling IBC based on M-Z modulation are discussed. Finally, in order to verify the proposed model, both the simulation and the *in vivo* measurements were carried out and the results were analyzed.

2 Modeling of IBC Based on M-Z Electro-Optic Modulation

2.1 Circuit Model

The electrostatic coupling intra-body communication based on M-Z electro-optic modulator can be described by Fig. 1, in which the signal electric field transmitting in the human body is input to the M-Z electro-optic modulator for optical phase modulation, and finally this change is converted into the amplitude change of electric signal by photodetector (PD). Due to the ground electrode is insulated from the receiver circuit, this kind of IBC receiver will result in low noise and low signal distortion.

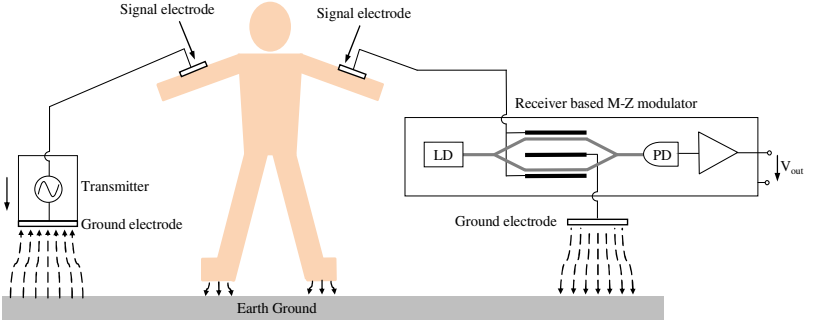


Fig. 1. The model of the electrostatic coupling IBC based on M-Z modulator

The electrostatic coupling IBC based on M-Z modulator can be described by the circuit model shown in Fig. 2. At the transmitting terminal of Fig. 2, IBC transmitter is considered as a voltage source with the output resistance of R_0. Meanwhile, due to the signal electrode is attached on the human skin, both the resistance and the capacitance between them should be considered. Considering the impedance of the return path is far greater than that of the body channel, the human body can be modeled as a perfect conductor [1], [7]. Therefore, the human body is considered as a body node. Meanwhile, the coupling capacitance between the human body and the earth ground is also represented as C_{g2}. At the receiving terminal, due to M-Z

modulator has two internal signal electrodes corresponding to two arms and one internal ground electrode, resulting in the two capacitances representing as C_{e1} and C_{e2}, as shown in Fig. 2. Additionally, the coupling capacitance between the ground electrode of the receiver and the earth ground is also represented as C_{g3}.

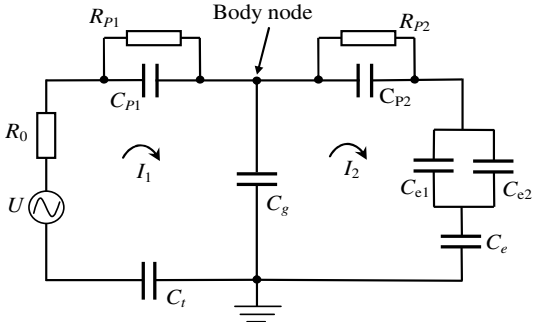


Fig. 2. The circuit model of the electrostatic coupling IBC based on M-Z modulator

2.2 Mathematical Model

Firstly, based on the proposed circuit model, the relationship between the output voltage (U_{in}) of IBC transmitter and the voltage applied on the M-Z modulator (U_e) can be achieved, as shown in the following

$$U_E = \frac{U_{in} Z_{g2} Z_e^{'}}{(R_0 + Z_{s1} + Z_{g1} + Z_{g2})(Z_{s2} + Z_e^{'} + Z_{g2} + Z_{g3}) - Z_{g2}^2} \cdot \tag{1}$$

Secondly, as for the transverse type EO modulation [2] shown in Fig. 3(a), the direction of the signal electric fields (z direction) applied on the x-cut EO crystal is perpendicular to the light propagation direction (y direction), while the phase delay of the output light can be expressed as

$$\Delta\phi = \frac{2\pi l}{\lambda_0}(n_z - n_x) = \frac{2\pi l}{\lambda_0}[(n_e - n_o) - \frac{1}{2}(n_e^3 \gamma_{33} - n_o^3 \gamma_{13}) E_z] \cdot \tag{2}$$

where l is the length of the EO crystal, and E_z represents the intensity of the electric field applied on the EO crystal.

As for the M-Z modulator used in IBC shown in Fig. 3(b), under the assumption that the M-Z modulator is also *x*-cut and *y*-propagating, the phase of arm A can be expressed as

$$\phi_a = \frac{2\pi l}{\lambda_0}(n_e - n_o) - \frac{\Gamma \pi l}{G \lambda_0}(n_e^3 \gamma_{33} - n_o^3 \gamma_{13}) U_e - \alpha_0 \cdot \tag{3}$$

On the other hand, due to the electric field direction of arm A is opposite to that of arm B, the light wave phase of arm A is also opposite to that of arm B. Therefore, the following equation can be derived.

$$P_{out} = P_{in} \exp[\frac{4\pi l}{\lambda_0}(n_e - n_o)] \cdot \cos^2[\frac{\Gamma \pi L}{G\lambda_0}(n_e^3 \gamma_{33} - n_o^3 \gamma_{13})U_e + \alpha_0] \quad \cdot \tag{4}$$

where P_{in} is the input optical power at the left 3dB coupler, and P_{out} is the output optical power at the right 3dB coupler shown in Fig. 3(b). Finally, the relationship between the output voltage (U_{in}) of the IBC transmitter and the detection voltage (U_{out}) of the IBC receiver based on M-Z modulator can be achieved.

$$U_{out} = \frac{SR_k}{10^{\frac{k}{10}}} P_{in} \exp[\frac{4\pi l}{\lambda_0}(n_e - n_o)] \cdot \cos^2\{\frac{\Gamma \pi L Z_{g2} Z_e^{'}(n_e^3 \gamma_{33} - n_o^3 \gamma_{13}) \cdot U_{in}}{G\lambda_0[(R_0 + Z_{s1} + Z_{g1} + Z_{g2})(Z_{s2} + Z_e^{'} + Z_{g2} + Z_{g3}) - Z_{g2}^2]} + \alpha_0\} \quad \cdot \tag{5}$$

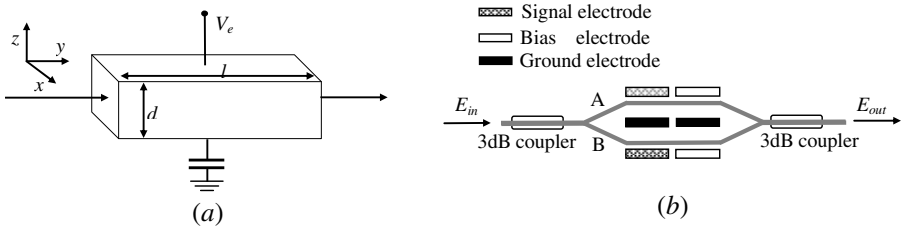


Fig. 3. The Electro-Optic modulator used in IBC. (a) Single EO modulator. (b) M-Z modulation.

2.3 Sensitivity Characteristic

Sensitivity is an important parameters of the IBC based on M-Z modulation, it determines the minimum threshold of the electoral signal injecting into the human body and the receiving voltage value corresponding to a transmitting voltage. In an IBC system, the minimum input optical power (ΔP_{out}) that can be converted into electrical signal by photodetector determines the minimum input voltage (ΔU_{in}) of IBC. Therefore, to achieve the sensitivity characteristic of the IBC based on M-Z modulation, the relationship between ΔP_{out} and ΔU_{in} should be discussed. Firstly, according to Eq. 4, the following equation can be achieved.

$$\frac{dP_{out}}{dU_E} = -\frac{P_{in}\Gamma \pi L(n_e^3 \gamma_{33} - n_o^3 \gamma_{13})}{G\lambda_0} \exp[\frac{4\pi l}{\lambda_0}(n_e - n_o)] \cdot \sin[\frac{2\Gamma \pi L}{G\lambda_0}(n_e^3 \gamma_{33} - n_o^3 \gamma_{13})U_e + 2\alpha_0] \quad \cdot \tag{6}$$

Generally, in order to achieve the linear relationship between U_e and P_{out}, the phase difference of α_0 should be set as $\alpha_0 = \pi/4$ by adjusting the voltage applied on the bias electrodes of M-Z modulator. As a result, when $U_e \to 0$, Eq. 6 can be simplified as

$$\frac{\Delta P_{out}}{\Delta U_e} = -\frac{P_{in}\Gamma \pi L(n_e^3 \gamma_{33} - n_o^3 \gamma_{13}) \exp[\frac{4\pi l}{\lambda_0}(n_e - n_o)]}{\sqrt{2}G\lambda_0} \quad \cdot \tag{7}$$

According to Eq. 7, we can found that P_{out} has an approximately linear relationship with U_e on condition that P_{in} keeps unchanged. Finally, based on the Eq. 1 and 7, the minimum input voltage (ΔU_{in}) corresponding to the minimum input power of the photodetector (ΔP_{out}) that can converted into electrical signal in an IBC system based on M-Z modulator can be expressed as

$$\Delta U_{in} = -\frac{[(R_0 + Z_{s1} + Z_{g1} + Z_{g2})(Z_{s2} + Z_E^{'} + Z_{g2} + Z_{g3}) - Z_{g2}^2]G\lambda_0 \cdot \Delta P_{out}}{\sqrt{2}Z_{g2}Z_e^{'}P_{in}\Gamma\pi L(n_e^3\gamma_{33} - n_o^3\gamma_{13})\exp[\frac{2\pi l}{\lambda_0}(n_e - n_o)]} \quad . \tag{8}$$

On the other hand, the relationship between ΔU_{out} and ΔU_{in} can also be achieved according to Eq. 8 and the optic-electro transfer characteristic of the photodetector.

$$\frac{\Delta U_{out}}{\Delta U_{in}} = -\frac{\sqrt{2}SR_kZ_{g2}Z_e^{'}P_{in}\Gamma\pi L(n_e^3\gamma_{33} - n_o^3\gamma_{13})\exp[\frac{2\pi l}{\lambda_0}(n_e - n_o)]}{10^{\frac{k}{10}}G\lambda_0[(R_0 + Z_{s1} + Z_{g1} + Z_{g2})(Z_{s2} + Z_E^{'} + Z_{g2} + Z_{g3}) - Z_{g2}^2]} \quad . \tag{9}$$

3 Experiments Results

To verify the validity of the proposed mathematical model and the sensitivity characteristic of the IBC based on M-Z modulation, both the simulation and the corresponding experiments based on the actual human body have been carried out.

3.1 Minimum Input Voltage

Firstly, the minimum input voltage (ΔU_{in}) corresponding to the minimum input power that can be converted into electrical signal by the photodetector can be determined according to Eq. 8. According to the parameters [1] provided by T. G. Zimmerman, the parameter of photodetector and the proposed model, the calculated result of ΔU_{in} is 0.672V. On the other hand, the measurement results of ΔU_{out} are shown in Fig. 4, in which the Fig. (a), (b), (c) and (d) are corresponding to U_{in}=0.1V, 0.2V, 0.5V and 1V respectively. We can found from fig. 4(a) that the receiver can't receive the signal corresponding to U_{in} = 0.1V. When U_{in} was increased to 0.2V, the receiving signal has severely distortion, as shown in Fig. 4(b), while it's difficult to use the receiving waveform corresponding to U_{in} = 0.2V as communication signal. Compared with Fig. 4(a) and (b), Fig. 4(c) corresponding to U_{in} = 0.5V has comparatively lower signal distortion. As U_{in}=1V, we can see from Fig. 4(d) that the receiving signal has acceptable waveform, which helps to achieve the data communication between two IBC devices. Therefore, according to the discussions mentioned above, the measurement results generally coincide with the simulation result based on Eq. 8 (ΔU_{in} = 0.672V).

3.2 Response

The simulation results of the relationship between U_{out} and U_{in} and the corresponding measurement results are shown in Fig. 5, which also provides the errors

corresponding to each measurement value of U_{in}. It can be seen from Fig. 5 that the mathematical simulation results based on Eq. 9 basically agree with the corresponding measurement results, while the absolute values of the errors between the simulation and the measurement are limited in 19.6mV. Moreover, both the simulation results and the *in vivo* measurement results approximately linearly increases as U_{in} increases from 0.2V to 5V, which indicates that there has an approximately linear relationship between U_{out} and U_{in} in the IBC system, while the proposed mathematical model has the approximately equal response with the IBC system based on the M-Z modulator.

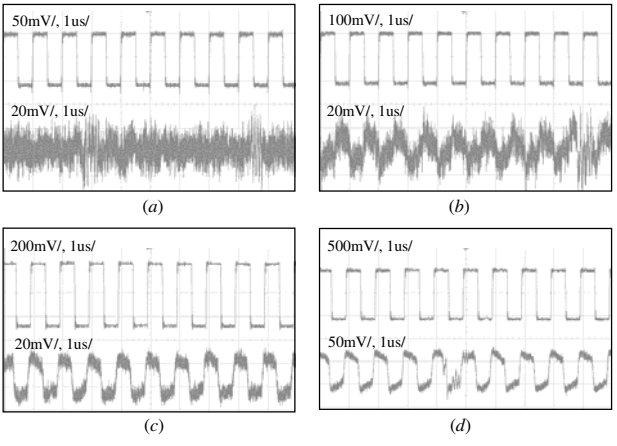


Fig. 4. Measurement results of ΔU_{out}. (a) U_{in}= 0.1*V*. (b) U_{in} = 0.2*V*. (c) U_{in} = 0.5*V*. (d) U_{in} = 1*V*.

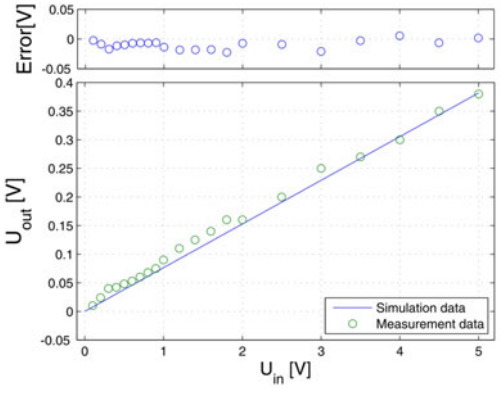


Fig. 5. The simulation and measurement results of the relationship between U_{out} and U_{in}

4 Conclusion

In this paper, the mathematic model of the electrostatic coupling intra-body communication based on M-Z optoelectronic modulation has been proposed, while

the sensitivity characteristics of this kind of IBC are discussed. Our investigation indicates that the simulation results of the sensitivity characteristic coincide with the corresponding *in vivo* measurement results, while the proposed model and sensitivity characteristic will help to achieve reliable, high speed and low cost intra-body communication.

Acknowledgment

The work is supported by National Natural Science Foundation of China 60801050. The authors are also grateful to the Institute of Optics, University of Rochester, America, for their help.

References

1. Zimmerman, T.G.: Personal Area Networks (PAN): Near-Field Intra-Body Communication. In: Media Art and Science. Master Thesis: Massachusetts Institute of Technology (1995)
2. Shinagawa, M., Fukumoto, M., Ochiai, K., Kyuragi, H.: A Near-Field-Sensing Transceiver for Intrabody Communication Based on the Electrooptic Effect. IEEE Trans. Instrum. Meas. 53(6), 1533–1538 (2004)
3. Wegmueller, M.S., Huclova, S., Froehlich, J., Oberle, M., Felber, N., Kuster, N., Fichtner, W.: Galvanic Coupling Enabling Wireless Implant Communications. IEEE Trans. Instrum. Meas. 58(8), 2618–2625 (2009)
4. Wegmueller, M.S., Oberle, M., Felber, N., Kuster, N., Fichtner, W.: Signal Transmission by Galvanic Coupling Through the Human Body. IEEE Trans. Instrum. Meas. 59(4), 963–969 (2010)
5. Sasaki, A.-I., Shinagawa, M., Ochiai, K.: Principles and Demonstration of Intrabody CommunicationWith a Sensitive Electrooptic Sensor. IEEE Trans. Instrum. Meas. 58(2), 457–466 (2009)
6. Sasaki, A.-I., Shinagawa, M.: Principle and Application of a Sensitive Handy Electrooptic Probe for Sub-100-MHz Frequency Range Signal Measurements. IEEE Trans. Instrum. Meas. 57(5), 1005–1013 (2008)
7. Zimmerman, T.G.: Personal Area Networks: Near-Field Intra-Body Communication. IBM Syst. J. 35(3-4), 609–617 (1996)

Analysis on the Starting Point and Path of Informatization in the Underdeveloped Rural Areas: A Survey of 52 Undeveloped Rural Villages of Ningbo*

Zhi-wei Yu, Guo-hong Zhou, and Bin Wang

Ningbo Institute of Technology, Zhejiang University, Ningbo, China
nityzw@gmail.com

Abstract. Accelerating the informatization in underdeveloped rural areas is significant for promoting the informatization in rural areas. Based on the current research, three fundamental questions are raised, namely, what are the need, the starting point and the path of informatization in the rural areas? According to the survey of 52 underdeveloped rural villages of Ningbo, the pattern of the starting point and the differentiated path are put forward to answer those three questions. Finally, some suggestions are given according to the state quo of the informatization in underdeveloped rural areas.

Keywords: the underdeveloped rural areas, informatization, the starting point, path, pattern, suggestion.

1 Introduction

The informatization of rural areas has become important to the construction of new socialist rural area. Since 2005, No. 1 Issue of the Central Committee has continuously focused on the informatization in rural areas. The informatization in rural areas is the important part of national economy informatization and the main power of the development of rural economy and the effective approach for the adjustment of industrial structure, reduction of market risk and increase of farmers' income[1].The government has paid much attention to not only from planning to implementation but also from policy to basic infrastructure. Some achievements have been gained. But in underdeveloped rural areas, it is relatively slow and is not optimistic.

From the perspective of the difficulties and exiting problems, Wang [2], Long[3] and Shao[4] contributed the situation to backward ideas, lack of fund, division of information in different departments, the quality of working staff, peasants' income, single means of information transferring, uneven distribution of information, lack of talents, unhealthy management system, hidden danger in information security and so on.

* This project were supported by the Philosophy and Social Science Planning Project of Ningbo(#G10-A46) and Zhejiang Provincial Natural Science Foundation of China(#Y6110693) and Zhejiang Provincial Education Science Planning Project of China(#SCG449).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 217–224, 2011.

Accordingly, the above scholars brought forward some countermeasures for the difficulties and exiting problems in underdeveloped rural areas as follows. The government should play the role of guide. The information service network should be improved. The construction of information teams and local information resources and integration should be enhanced. Long run operation management mechanism should be established and information service system should be strengthened. In addition, Lan[5] appealed to application of information network in the whole underdeveloped rural areas.

Informatization in different regions of rural areas is unbalanced. The underdeveloped areas lag behind the developed areas greatly. So it is a long way to go for the informatization in underdeveloped rural areas[3]. How to combine the actual facts in underdeveloped rural areas is the entry and path to research the informatization in the underdeveloped rural areas, which has the realistic meaning for acceleration of the informatization in underdeveloped rural areas and is helpful for the solution of "*the last mile*" and narrowing the "*digital gap*".

2 The Status Quo of Informatization in the Underdeveloped Rural Areas

2.1 Research Objects and Questionnaire Design

Several typical problems should be solved if we research the starting point and path of the informatization in underdeveloped rural areas. *Firstly,* what is the demand of the informatization in the underdeveloped rural areas? *The second* is how to locate the starting point of the informatization construction in underdeveloped rural areas. *The third* is how to choose the developing path.

The informatization in the underdeveloped rural areas is a complicated project and a long process of constant building and improvement. The current studies have not indicated the clear path to construct it.

Questionnaires designed according to the relevant policy were made by investigating 727 farmers (about 15 farmers each village at random selection) after the interview of 52 villagers committees.

The content was designed with reference to valuable experience and successful applications from such developed countries as American, Japan, French and German to other countries such as Korean, India and Indonesia[6,7]. The topics were about the information infrastructure, information awareness, demand, and technology application and information resource construction[8].

2.2 The Status Quo of the Unerderdeveloped Rural Areas

According to statistical analysis of Ningbo, the status quo of the informatization in the underdeveloped rural areas is mainly shown as follows.

(1) The informatization infrastructure of rural areas is rather good. With the government's driving forward, the national comprehensive information service platform about agriculture has developed greatly. With the great support of Ningbo Municipal Government, Ningbo Information Industry Bureau and other relevant

government departments, the information infrastructure in the underdeveloped rural areas has made certain achievement in recent years. The improvement of network year after year has made a good foundation. The ownership of traditional information equipment such as TVs, telephones and cell phones has reached 89%, 83% and 79% respectively. Most people in charge of village committees reflected that more and more villagers had bought cell phones, computers and so on under the guide of the information service station in the rural areas. The information equipment keeps a steady growth. Although the percent of computer ownership is fairly low, 80% of the households who have computer have opened the broadband.

(2) The informatization awareness is generally weak. Because of the literacy constraint, economic condition, geography and other factors, farmers in the underdeveloped rural areas don't know much about the network, which makes farmers to know little about the informatization in the rural areas and make little use of the information resources in their production and daily life. The survey showed that nearly 40% of the farmers never used the internet and time spent on the internet was also rather short. Only 6% of the farmers thought it was helpful to use internet. Most of them couldn't make sure whether network was helpful. The negative effect of network caused misunderstanding for those that never touched the internet. Most farmers showed ignorance about online trade, the remote training and other forms of informatization. Most people thought it was only an entertaining tool. Only 25% of the farmers thought they could know the information and policies about agriculture and rural areas.

(3) The hierarchy of information need is not high. For permanent residents in the countryside, they didn't have strong desire to obtain information. So the hierarchy of information need was not high. More than 30% of the villagers didn't know how to get information. The demand was only limited on the knowledge about daily life, weather forecast and so on through such traditional tools as TV, the bulletin board and leaflets.

(4) The informatization facilities don't play the role as expected. The informatization quality of the cadres in charge of the village committees was low. Lack of professional skills and low profession level led to the fact that there was only hardware but no software. Some villagers knew nothing about the role of the information service station and even the existence. 70% of the villagers have never been to the information service station of the village. And the information service stations have not worked. The infrastructure was divorce from the information service greatly. For the family who had the equipment, it is mainly used to read news, watch movies and other entertaining activities. Only 18% of the villagers used the tool to trade. 31% of the villagers used to search for information and 26% used to learn knowledge. Although the informatization infrastructure in the underdeveloped rural areas has been improved year after year, the usage of the new facilities such as computers, cell phones needs exploring. The equipment utilization rate was very low and the equipment wasn't made full use of.

(5) Lack of acquaintance of information about agriculture, farmers knows little about the current information resource platforms. They just hear of one or two platforms occasionally, such as *www.cnluye.com*. They know little about such mature information systems as *www.nb110.com, www.zjnm.com*. These systems are approved by the agriculture experts of Ningbo. On the other hand, the villagers also knew little

about the direct and related agencies such as Ningbo Municipal Bureau of Agriculture, Ningbo Municipal Bureau of Water Conservancy and Ningbo Municipal Bureau of Agriculture Machinery. What caused the situation was not only farmers' lack of acquaintance of information resources but the weak advertisement of the villager committee. As a result, there was lack of effective communication among farmers, the villager committee and the senior government department in charge of agriculture.

3 The Starting Point Analysis of Informatization Construction in the Underdeveloped Rural Areas

To choose the right starting point is the key for accelerating the general informatization. Informatization is a long-term systematical project and promises no quick results. The informatization needs to combine the local situation closely and uses the information technology and the website to help the farmer to solve the urgent problems at present to enable the farmers to recognize the value of informatization as soon as possible. Therefore, the starting point should be found out according to the actual situation in the rural areas.

(1) Choosing the starting point according to the infrastructure difference. As shown in table 1, the ownership rate of computers is not high. Many farmers even don't have the experience of using computer and surfing the internet. According to the questionnaire towards the farmers, 39.2% of the farmers have never used the internet. As for the areas with the poor infrastructure, the main task is to further strengthen the informatization infrastructure. The infrastructure is the basic construction of informatization in the underdeveloped rural areas.

Thus, for the rural areas with imperfect infrastructure, the government should give strong support and further regulate and enhance the construction of the information service station in the rural areas to make the farmers use computers and the network freely so as to recognize the value of the computer and the internet gradually.

(2) Choosing the starting point according to the difference of the information resource in the rural areas. According to the questionnaires done by the writer, the information service stations are not made full use of. The main reason is lack of the knowing the farmers' need and the guidance to farmers. From the survey, we can know that the urgent information need at present for farmers in the underdeveloped rural areas is daily life information. It concludes the general information as the general knowledge and agriculture news the main content, the specialized information such as meteorological information, planting technology, culture technology and market information such as the price, supply and need. In addition, farmers in the underdeveloped rural areas need practical, easy and affordable information service. But the advertisement of the public information platform is not enough. And many farmers lack understanding. On the other hand, the existing information doesn't make the difference. Therefore, when we promote the information platform, we should be around the characteristics in the rural areas and the urgent information demand for the farmers to conduct the information resource construction as so to guide and stimulate farmers to take part in the informatization in the rural areas.

Table 1. A survey about computers and broadband usage in the underdeveloped rural areas

County (City)	Township (Town)	The number of the villages investigated	Average computer ownership rate(%)	Average broadband opening rate(%)
Ninghai	Shenzhen Town	7	7.25%	89.10%
	Chayuan County	6	4.02%	59.20%
	Sangzhou Town	4	3.96%	53.03%
	Chalu Town	3	13.17%	86.56%
	Qiantong Town	4	36.51%	94.77%
Yuyao	Dalan Town	5	20.78%	91.05%
	Liangnong Town	5	34.35%	100.00%
Fenghua	Xikou Town	7	17.36%	99.25%
Xiangshan	Hepu Town	3	6.82%	63.49%
	Xinqiao Town	8	14.60%	97.47%

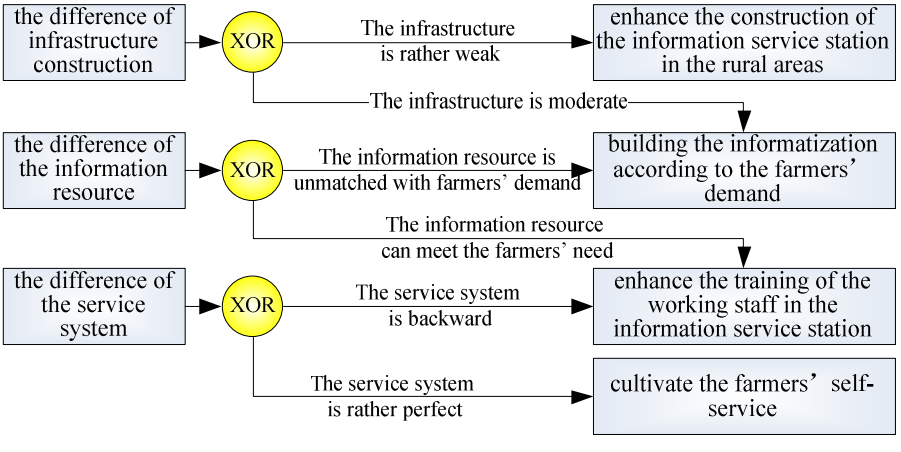


Fig. 1. The starting point pattern of the informatization in the underdeveloped rural areas

(3) Choosing the starting point according to the difference of rural information service system. The workers of the information service stations are the bridge which can shorten the distance between the farmers and the information service in the rural areas. The survey indicated that the number of the workers was small. There were one or two workers in most information service stations and their daily service has little collection with the informatization. The advertisement of the information platforms

about agriculture and the further understanding of the farmers' demand needs to be enhanced. When the information resource construction can meet the farmers' need, the workers in the information station can play a better guiding role and teach the farmers themselves to make good use of the information stations and promote their information application ability so as to make the computers and the network play the positive role in the economy, politics and culture. In this way can effectively promote the informatization and social harmony in the rural areas.

To sum up the above ideas, the starting point pattern of the informatization is made from the differences of the informatization, the information resource construction and the service system according to the differences of the informatization in the rural areas on the basis of the analysis of the characteristics of the information need(as in Fig. 1).

4 Path Analysis of the Informatization in the Underdeveloped Rural Areas

Combining the status quo of the farmers and rural areas, the feasible informatization path can be explored (as shown in Fig. 2) according to the differences of the informatization infrastructure construction, information resource construction and application, the information service system and the construction of the information team.

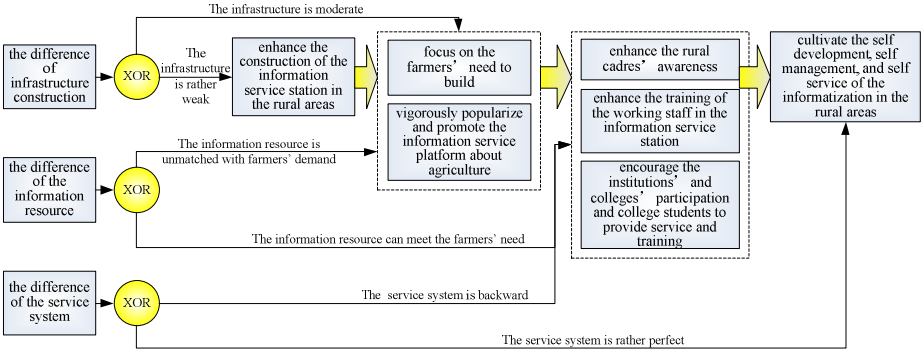


Fig. 2. The difference of the informatization path in the underdeveloped rural areas

(1)Strengthening and regulating the construction of the information service station in the rural areas. For the counties and towns with weak informatization infrastructure, the construction of the information service station needs further strengthening and regulating so as to draw the farmers' attention to know, consult and use the information they need. In this way do the farmers promote the informatization awareness and complete the shift from the informatization infrastructure as the main emphasis to the farmers' own informatization facilities.

(2)Vigorous publicizing and promoting of existing information services platform about agriculture. The work of vigorous advocating and popularizing the information

service platform about agriculture needs to be done. And the farmers should be organized orderly and trained to know the general knowledge, agricultural news, weather, planting technology and cultivation technology and so on. It will stimulate the farmers to make use of the information service station and the information platform initiatively to enable the rural areas with the better informatization infrastructure to improve the equipment utilization rate and promote the value of the information facilities so that the farmers can further know and use the present information resources.

*(3)Developping and improving the information service platforms about agriculture with the farmer's deman*d. Concerning the farmers' demand, the information service platform about agriculture is developed and improved. A comprehensive research should be conducted to get the information need of agriculture, rural areas and farmers. The methods of the informatization should be used to meet the farmers' urgent need. The e-commerce platforms about agriculture should be developed and improved with the special characteristics in the rural areas.

(4)Improving the sense of informatization of the rural cadres and enhancing the training of working staff in the information service stations. Farmers should be guided to make use of the relevant information resource step by step with the help of the rural cadres and the information workers and improve their information quality so as to make the best use of the positive effect of the computer and the internet in the economic, social and cultural fields. In this way can the informatization and the social harmony be promoted rapidly.

(5)Cultivating the self development, self management and self service of the informatization of farmers and rural areas. Through construction and improvement of the informatization infrastructure and the popularization of the information resource and characteristic building and the improvement of quality, farmers can be guided to use the information in the daily life and agricultural production and sale and e-commerce of the specific resources spontaneously. So the farmers can feel the value of informatization and then the goal of self development, self management and self service can be achieved and the problems of "the last mile"and the digital gap can be solved.

5 Suggestions on Vigorously Pushing on the Informatization in the Underdeveloped Rural Areas

Through the analysis of the status quo, the starting point and the path of rural informatization, four aspects need to be done well in the process of pushing on the informatization in the underdeveloped rural areas.

(1) Further strengthening the government's main guidance role in the informatization in the underdeveloped rural areas. The informatization infrastructure is further improved by the financial support. The government encourages the institutions and colleges and enterprises to take part in the informatization in the underdeveloped rural areas and perfects the information service platform about agriculture to meet the need in the rural areas. At the same time the promotion and publicity of the information service platform can be done by the ownership of IVs and telephones in most families. Besides, the college graduates should be encouraged to serve the informatization in the underdeveloped rural areas.

(2) Vigorously improving rural cadres' informatization awareness. The rural cadres are the backbone in the informatization. Various activities such as effective training and learning can be organized and the corresponding check-up system should be set up to cultivate the information awareness and improve the rural cadres' initiative and enthusiasm.

(3) Making great efforts to train the working staff in the information service station. With the talent resource from the colleges and enterprises, the professional training should be conducted for the working staff in the information service stations to improve their profession standard and service awareness, and to create the service methods so as to make full use of the building and promoting role in the informatization.

(4) Combining the rural situation and meeting the farmers' practical need by the information method quickly. Farmers' extensive participation in the informatization is the most important task and the goal in the informatization. Informatization is a long process, which needs to use the innovative methods, combine the rural specialized resources and demands, explore the corresponding information resources and service, solve the farmers' urgent problems, and stimulate the farmers' extensive participation so that the beautiful prospect of self development, self management and self service can be achieved at last.

References

1. Li, Z.-S., Yang, S., Zhang, H., et al.: Establishment and Improvement of Agricultural Information System in China. Journal of Anhui Agricultural Sciences 34(14), 3544–3546 (2006)
2. Wang, X.: Research on the Strategy of Informatization Development in Underdeveloped Areas. A Dissertation Submitted to Tianjing University for the Degree of Marster (December 2007)
3. Long, C.: Information Construction Strategy in Less-developed Rural Mountain Areas. Journal of Intelligence 2, 187–189 (2009)
4. Shao, K.-D.: Research on the Countermeasures for Higher Vocational College Serving the "Three Agricultural" Informatization in Undeveloped Areas. Journal of Anhui Agricultural Sciences 37(22), 10825–10828 (2009)
5. Lan, J.: Zhejiang: speed up informatization in underdeveloped areas. China Information Times 20, 26 (2004)
6. An, X.: The Ideas and Countermeasures of Agricultural Informatization in China. Rural Economy 11, 90–91 (2005)
7. Wang, P.: On Agricultural informatization in Liaoning Province. Liaoning Agricultural Sciences (5), 38–39 (2006)
8. Yang, C., Jiang, Z.: Constructs the Evaluating Indicator System for China's Countryside Informationization. Journal of Intelligence (2), 24–27 (2009)

Multi-view Models for Cost Constitution of Cloud Service in Cloud Manufacturing System

Y. Cheng[*], D. Zhao[2,**], A.R. Hu[1], Y.L. Luo[1], F. Tao[1], and L. Zhang[1]

[1] School of Automation Science and Electrical Engineering,
Beihang University, 100191 Beijing, China
[2] Department of Electrical and Compter Engineering,
The University of Michigan-Dearbom, MI 48128-1491, USA
dmzhao@umich.edu

Abstract. Cloud manufacturing (CMfg) is a new manufacturing paradigm proposed recently. The existing research on CMfg stays on its concept, architecture, and running model. One of the typical characters of CMflg is on demand use and pay of manufacturing capability and manufacturing. Therefore, it is essential to investigate the cost issues of manufacturing cloud service in CMfg. The cost of cloud service (CS) in CMfg is studied from three different aspects in this paper, and multi-view models of cost are proposed, they are (a)The first phase is to build the cost constitute model viewed on the whole life-cycle of CS in CMfg. The second phase is the cost constitute model from the view point of related manufacturing enterprises in which providing, composing, and invocating cloud services are provided. Finally, the cost constitute model based on the cloud manufacturing services platform is evaluated.

Keywords: Cloud manufacturing (CMfg); Cloud service (CS); Cost; Multi-View model.

1 Introduction

Cloud manufacturing (CMfg) is a new service-oriented manufacturing model, which is developed from existing advanced manufacturing modelsunder the support of cloud computing technologies, internet of thing, high performance computing technologies, virtualization and service-oriented technologies, etc. [1-4]. At present, the research on CMfg is in its initial stage, and most of research topics are about its concept [1], architecture of system and platform [1], the running models [3-4], and the core enabling technologies [1-3]. One of the typical characters of CMflg is on demand use and pay of manufacturing capability and manufacturing. Therefore, it is necessary to study business operations of CMfg in order to implement the principles and concepts in practice. Cost constitution is one of the key issues that determine practical business operations of CMfg and achieve the lowest cost target.

[*] Y Cheng, master student of Beihang University (Beijing University of Aeronautics and Astronautics). The main research direction is the service comprehensive utility evaluation in service-oriented manufacturing system.
[**] Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 225–233, 2011.

In the authors' previous work [5], resource service transaction (RST) has been investigated to achieve the maximal comprehensive utility with regard to revenue, time and reliability. But the the cost constitution is not analyzed in detail. Some other related works of cost analysis of services can be found in the fields of grid computing [6], cloud computing [7], software engineering [8], and advance manufacturing systems[9]. However, in order to realize the lowest cost of the whole CMfg system, one has to consider the detailed costs of software, hardware, and other materials, the detailed costs of each manufacturing enterprises, and cloud manufacturing service platform. Therefore, in this paper, the multi-view model for cost constitution of cloud service (CS) is proposed based on the whole life-cycle of CS, the related manufacturing CS enterprises, and the cloud manufacturing service platform.

2 Description of Cloud Service

Manufacturing resource in CMfg is the general concept of various software and hardware resources, human resources, knowledge, etc. It is composed of manufacturing resource and manufacturing capability [2, 10], including hard manufacturing resource (i.e., machine tools, computing equipments, simulating equipments, testing equipments, materials, and other various manufacturing hard equipments), soft manufacturing resource (i.e., various models, data, software, information, knowledge in manufacturing process, etc.), and manufacturing capability (i.e., design, product, simulation, maintenance, management capability, and other capability in manufacturing process).

According to the running model of CMfg [3], there are three kinds of users in CMfg system, i.e., manufacturing cloud service provider (CSP), the operator or agent of manufacturing cloud service (C-Agent), and manufacturing cloud service demander (CSD), which providing, composing, and invocating cloud services, and realizing cloud service transaction. In the activities of CS transaction, the utility equilibrium can help each user obtain their own maximal utility and be useful to make the maximal utility of the whole system. However, cost is one of the most important issues of the utility, and it is necessary to investigate the cost constitution of CS.

3 Cost Constitution Model Viewed on the Whole Life-Cycle of CS

According to [2], the whole life-cycle process of CS is illustrated in Fig. 1., which is divided into the following steps:

(1) Manufacturing cloud service realization, i.e., manufacturing resource and manufacturing capability are virtualized and encapsulated into manufacturing cloud service based on knowledge.
(2) Cloud service aggregation and generating manufacturing cloud.
(3) Cloud service invocation.
(4) Cloud service decomposition, or generating new manufacturing resource and manufacturing capability.

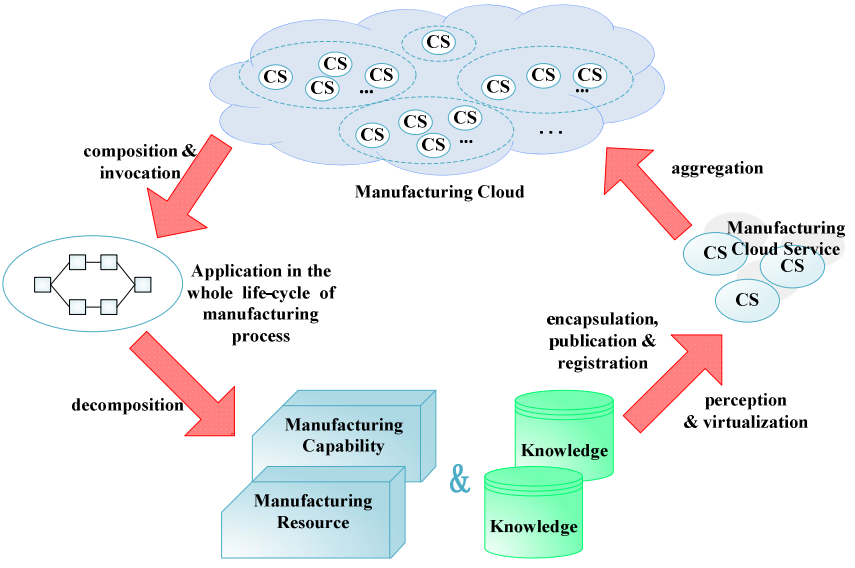


Fig. 1. The whole life-cycle of cloud service

Cost constitution of CS's application in the whole manufacturing life-cycle process

		prophase		metaphase		anaphase	
		programming	design	simulation & production	assembly & testing	logistics & distribution	running, after sales service & recovery
Management cost class	Human resource cost	management personnel's wage and various welfare					
		design, research and plan personnel's wage and various welfare		produce and assembly personnel's wage and various welfare		logistics and distribution personnel's wage and various welfare	after sales service and maintenance personnel's wage and various welfare
	Material resource cost	office supplies purchase cost, material depreciation and maintenance cost, water and electricity fee, network and Internet cost…					
	Other resource cost	office travel cost, union activity funds, business treat cost, staff education and training funds, various tax…					
Operation cost class		planning consultation cost, market analysis and investigation cost, design and computing analysis cost, cooperative research cost, software updating and maintenance cost, business outsourcing cost, law, IT, knowledge service cost …		purchase transportation cost, tax and commodity inspection cost, Equipment updating and maintenance cost, software updating and maintenance cost, machine and material consumption, work shut-down loss, business outsourcing cost …		financial insurance, leasing cost, e-business cost, advertisement cost, packaging cost, logistics transportation cost, handling cost …	installation and debugging cost, running maintenance cost, recovery cost, logistics transportation cost, handling cost …

Fig. 2. The cost constitution of cloud service invocation

The cost constitution of the whole life-cycle of CS can be investigated as follows:

(1) *Cloud service realization cost*, including resource perception cost, resource connection cost, resource virtualization cost, and resource encapsulation cost, etc.

(2) *Cloud service aggregation cost*, including computing cost, storage device cost, and CS management cost, etc.

(3) *Cloud service invocation cost*. It primarily reflects the application of CS in the whole life-cycle of manufacturing process. Cloud service invocation cost can be divided into management cost class and operation cost class, as shown in Fig. 2.

a) *Management cost class* consists of human resource cost, material resource cost, and other resource cost.

Human resource cost includes management personnel, design personnel, research personnel, plan personnel, produce personnel, assembly personnel, logistics personnel, maintenance personnel and all other personnel's wage and various welfare, etc.

Material resource cost involves office supplies purchase and replace cost, material depreciation cost, material maintenance cost, water, electricity, and other power cost, local network, Internet and other communication cost, etc.

Other resource cost covers office travel cost, union activity funds, business treat cost, staff training funds, various tax, etc.

b) *Operation cost class* contains all costs related to the business operation activities, i.e., consultation cost, market analysis and investigation cost, business outsourcing cost, law operation cost, purchase cost, logistics transportation cost, handling cost, tax and commodity inspection cost, financial insurance, etc.

(4) *Cloud service decomposition cost*, including the new generated manufacturing resource and manufacturing capability storage cost, decomposed CS maintenance cost, etc.

Besides the classification based on the above steps in the whole life-cycle, the whole life-cycle cost constitute of CS can also be divided into hardware cost class, software cost class, and other cost class, which is similar to the thought of the cost constitute analysis based on cloud manufacturing service platform in next section.

4 Cost Constitution Model Viewed on the Related Manufacturing CS Enterprises

The related manufacturing CS users are CSP, CSD and C-Agent. One manufacturing enterprises can be both CSP and CSD. Several parameters as follows have to be taken into account in order to estimate the cost constitution of CSP and CSD, as shown in Fig. 3.

(1) *Business premises cost*, including buildings cost, transportation tools cost, electricity, water, and other power cost, etc.

(2) *Hardware cost*, including new hardware resources (i.e., machine tools, computers, materials, et al.) purchase cost, hardware components replacement cost, hardware maintenance cost, etc.

The total costs of manufacturing cloud service provide and demand enterprises	**Business premises cost**	buildings, electricity, transportation tools...
	Hardware cost	new resources（machine tools, computers, materials...）
		replacement of hardware components
	Software cost	CAD/CAXA/Master CAM/Solidworks/CATIA/Matlab...
	Personnel cost	management staff
		mechanic, maintenance personnel
		experts, technology developers or researchers
	Knowledge cost	design manual, standard, protocol...
		experience and capability to design, product, simulating testing, maintain, manage...
	Communication cost	connection to the Internet（including cost for using the CMfg service platform）
		local area networks
		telephone/fax networks
	Logistics cost	consumer service, demand forecast, order management, transportation, storage management, portage and handling, purchase, package...

Fig. 3. The cost constitution of CSP and CSD

(3) *Software cost*, including all kinds of software (i.e., design software, computing software, simulation software, and other software) purchase cost, developing cost, updating cost, and maintenance cost, etc.

(4) *Personnel cost*, classifying into three kinds as management staff cost, mechanic and maintenance personnel cost, and experts, technology developer or researcher cost, and other service personnel cost according to the staff classification. It also divides into personnel's wage cost and personnel's welfare two types.

(5) *Knowledge cost*, including the costs of design manual, standard, protocol, and other tangible knowledge, and the costs of experience, capability, and other intangible knowledge to design, product, simulating, testing, management, and other manufacturing activities.

(6) *Communication cost*, including the cost of connection to the Internet, local area networks cost, and telephone and fax networks cost. As to CSP and CSD, the cost for using cloud manufacturing service platform is considered into the cost of connection to the Internet.

(7) *Logistics cost*, the important cost for distributed manufacturing resources. It involves consumer service cost, demand forecast cost, order management cost, transportation cost, storage management cost, portaging and handling cost, purchase cost, package cost, etc.

As to C-Agent, the most important operation and management function is depended on cloud manufacturing service platform. Besides the above cost constitution of CSP and CSD, C-Agent still has the cost to build and maintain cloud manufacturing service platform which analyzed in next section in detail.

5 Cost Constitution Model Viewed on Cloud Manufacturing Service Platform

According to manufacturing cloud construction levels [10], the set-up process of cloud manufacturing service platform consists of six layers, i.e., resource layer, resource perception layer, resource virtualized connection layer, manufacturing cloud core service layer, transmission network layer, and terminal application layer. On the basis of those six layers, the cost constitution can be analyzed as hardware costs, software costs, and other costs with each layer. As shown in Fig. 4., the cost constitution model of CS based on cloud manufacturing service platform is made up of those six layers costs from-bottom-to-top as follows:

(1) Cost of physical resource layer (P-Layer)

P-Layer makes various physical resources be accessed to network and realizes the overall interconnection of physical resources, and then forms the virtual resource to provide the interface for virtual resource encapsulation and invoking. In P-Layer, various manufacturing hardware resource purchase, renewal and maintenance cost, various manufacturing software resource purchase, updating, and maintenance cost, and various manufacturing capability generation cost are considered to analyze the cost constitution, as well as human resource cost, power cost, supporting knowledge and technology cost, and other costs. Most of the costs in R-Layer are assumed by CSP and CSD.

(2) Cost of resource perception and connection layer (C-Layer)

C-Layer makes distributed resource be perceived and connected into CMfg service platform, and forms virtual resource pool. Hardware needed in the perception and connection process includes identification equipments, adapters, and sensing equipments, etc. The hardware cost of C-Layer contains identification equipments, adapters, and sensing equipments purchase, renewal, and maintenance costs. The software cost of C-Layer is the cost to develop, update, and maintain resource perception system information center. Personnel cost, electricity, water and other power cost should be taken into account.

(3) Cost of virtual resource layer (R-Layer)

The primary function of R-Layer includes cloud service definition, cloud service publish management, resource quality management, resource provider pricing and accounting management, and resource division management, etc. Therefore, the

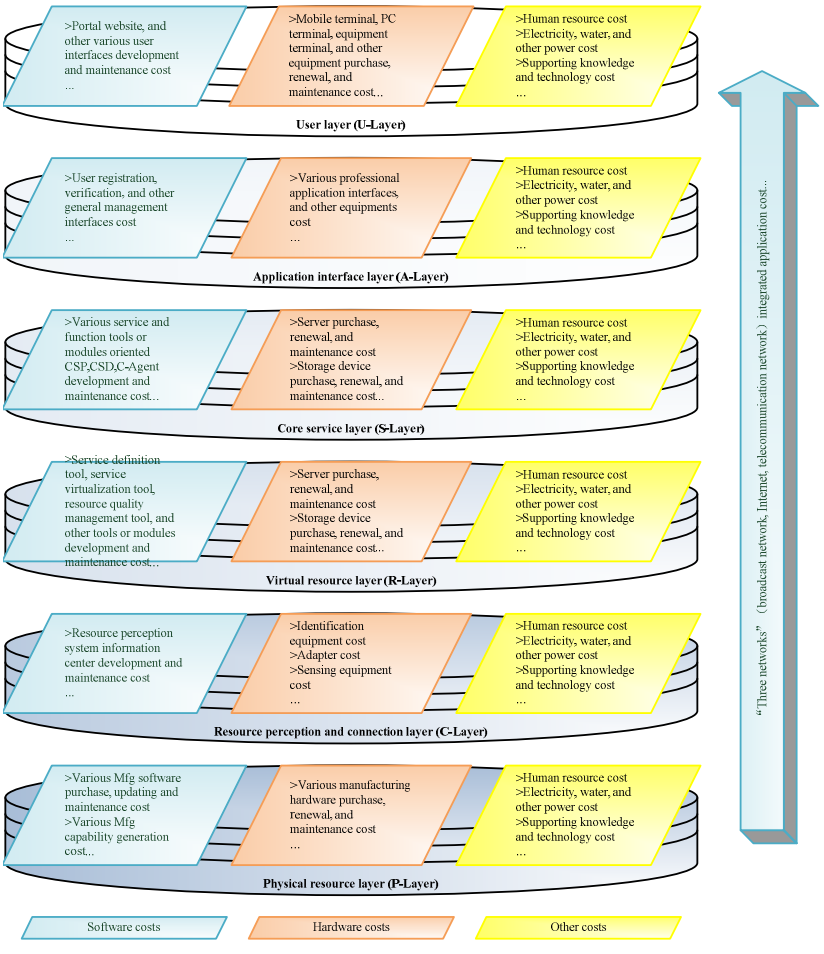


Fig. 4. Cost constitution model view on cloud manufacturing service platform

expenses is composed of the costs to develop and maintain cloud service definition tool, cloud service publish management tool, and other various tools and modules. Server cost, storage device cost, and other hardware cost are also indispensable.

(4) Cost of core service layer (S-Layer)

S-Layer provides various core services and functions of manufacturing cloud service management for the three users (i.e., CSP, CSD, C-Agent). The costs of S-Layer contains the cost of CS standardization and testing management tool, and interface management tool, and other tools oriented to CSP, the cost of cloud task management tool, high-performance search tool, and other tools oriented to CSD, and the cost of user management tool, system management tool, CS management tool, data management tool, and other tools oriented to C-Agent. It also needs to pay attention to the cost of servers, storage devices, and other equipments' purchase, renewal and maintenance.

(5) Cost of application interface layer (A-Layer)
In A-Layer, different professional application interfaces and general management interfaces are provided. Therefore, it has to cost much to purchase, updating or renewal, and maintain various professional application interfaces and general management interfaces.
(6) Cost of user layer (U-Layer)
Various user interfaces and CS portal website are provided for various manufacturing fields and industries in U-Layer. Therefore, the software cost of U-Layer includes portal website, and other various user interfaces development and maintenance cost. The hardware cost of U-Layer is mainly used to purchase, replace and maintain mobile terminals, PC terminals, equipment terminals, and other equipments.

From bottom to top, those six layers still has the third cost class, i.e., the other cost, other than the hardware cost class and the software cost class. The six layers are connected by broadcast network, the Internet and telecommunication network. These three networks integrated application cost is also taken into account necessarily.

6 Conclusion

The study on the cost constitution of cloud manufacturing service is one of the key issues that determine the practical business operation of cloud manufacturing and realize the lowest Cost goal. With respect to the cost constitute of CS, the multi-view models are proposed in this paper, including (a) the cost constitute model based on the views from the whole life-cycle of CS in cloud manufacturing, (b) the cost constitute model based on the views from the related manufacturing enterprises which provide, compose, and invocate services, and (3) the cost constitute model based on the cloud manufacturing service platform.

This work is currently conducted as basic research on the business operation of CMfg. As one of the comprehensive utility evaluation indexes, it is necessary to further quantitatively investigate the cost of CS in order to obtain the mathematic models of the cost constitute, and then to realize the maximal comprehensive utility of each enterprise and the whole system.

Acknowledgments

This work was financially supported by the NSFC projects (No.51005012 and No.61074144) and the Fundamental Research Funds for the Central Universities in China.

References

1. Li, B.H., Zhang, L., Wang, S.L., Tao, F., et al.: Cloud manufacturing: a new service-oriented networked manufacturing model. Computer Integrated Manufacturing Systems 16(1), 1–7 (2010) (in Chinese)
2. Tao, F., Zhang, L., Venkatesh, V.C., et al.: Cloud manufacturing: a service and computing-oriented manufacturing model. Journal of Engineering Manufacturing (to be printed 2011)

3. Tao, F., Cheng, Y., Zhang, L., et al.: Cloud manufacturing. In: Advanced Materials Research, vol. 201-203, pp. 672–676 (2011)
4. Luo, Y.L., Zhang, L., He, D.J., et al.: Study on multi-view model for cloud manufacturing. In: Advanced Materials Research, vol. 201-203, pp. 685–688 (2011)
5. Cheng, Y., Tao, F., et al.: Study on the utility model and utility equilibrium of resource service transaction in cloud manufacturing. In: IEEE International Conference on Industrial Engineering and Engineering Management (IEEM 2010), Macau, December 7-10, pp. 2298–2302 (2010)
6. Alek, O., Hartmut, K., et al.: What does grid computing cost? Journal of Grid Computing 6(4), 385–397 (2008)
7. Liu, J.: Cost utility research of cloud computing. The Excellent Master Degree Thesis in China, Beijing University of Posts and Telecommunications (2010) (in Chinse)
8. Li, B.H., Zhang, L.: Comparison and choice of typical software cost estimation methods. Computer Integrated Manufacturing Systems 14(7), 1441–1448 (2008) (in Chinse)
9. Zhang, M.M.: Research on CIMS-based cost management models. The Doctor Degree Thesis in China, Xiamen University (2006) (in Chinse)
10. Zhang, L., Luo, Y.L., et al.: Key technologies for the construction of manufacturing cloud. Computer Integrated Manufacturing Systems 16(11), 2510–2520 (2010) (in Chinse)

Modeling and Analysis of Player Efficiency Rating for Different Positions: Case Study with NBA

Tong Zhang, Junlong Chen, and Xinchao Zhao

School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract. By means of data statistics, comparative analysis and mathematical modeling, the value of NBA player efficiency rating has been modeled and analyzed. The optimization model of the NBA player efficiency rating at different positions is correspondingly obtained. Based on the analysis of the typical players the optimization model can effectively reflect the technical needs of different locations in the match and is more targeted than the NBA's official formula of player efficiency rating computation.

Keywords: Player efficiency rating, NBA, optimization model, correlation coefficient, different positions.

1 Introduction

There are five positions in the modern basketball match: Power Forward (PF), Small Forward (SF), Center (C), Point Guard (PG) and Shooting Guard (SG)[1]. Players in different positions have different responsibility and different playing skills. Therefore, while computing the efficiency rating, such difference should be more considered. Obviously, the best male basketball players are gathering in NBA, which is generally accepted as the male basketball competition of the highest level. So doing research on those players is more representative and more valuable. The research on the efficiency rating of players could also be a beneficial reference for the association and the teams as well as the players themselves.

2 Objects and Methods of Research

2.1 Objects of Research

We use the technical statistics such as points, backboard, steals, assists, etc. and the official player efficiency rating (PER) of 306 players belonging to 30 teams in 2010-2011 season [2](until March 1 and excluding those whose rates of attendance are below 60%) as the original references.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 234–242, 2011.

2.2 Methods of Research

2.2.1 Research on Literature References

We search the hoopchina.com[3], official site of NBA China[4] and the site of ESPN[5] to get the technical statistics and the official player efficiency rating[2] of 306 players belonging to 30 teams in 2010-2011 season.

2.2.2 Mathematical Statistics

We induce those statistics, including field goal attempts、field goal, 3-point goal, free throw attempts, free throw, offensive rebound, defensive rebound, assist, personal foul, steal, turnover and block shot. In order to make our results more convincing and realistic, we are trying to utilize more research factors, analyzing the statistics by Excel 2007[6][7] and SPSS 13.0[8][9].

2.2.3 Analytic Method of Comparison

Analyze and compare the references and statistics that we get..

3 The Official PER Formula of NBA

NBA uses official player efficiency rating to evaluate the performance of players. All calculations begin with what is called unadjusted PER (uPER).The formula is:

$$\begin{aligned} uPER = {} & (1/MP) \times \{3P + (2/3) \times AST + (2 - factor \times tmAST/tmFG) \times FG + FT \times \\ & 0.5 \times [1 + (1 - tmAST/tmFG)] + (2/3) \times tmAST/tmFG - VOP \times TO - VOP \times \\ & DRBP \times (FGA - FG) - VOP \times 0.44 \times [0.44 + (0.56 \times DRBP)] \times (FTA - FT) + \\ & VOP \times (1 - DRBP) \times (TRB - ORB) + VOP \times DRBP \times ORB + VOP \times STL + \\ & VOP \times DRBP \times BLK - (PF \times lgFT/lgPF) - 0.44 \times lgFTA/lgPF \times VOP \}. \end{aligned} \tag{1}$$

Where:

$$factor = (2/3) - (0.5 \times (lgAST / lgFG)) / (2 \times (lgFG / lgFT)). \tag{2}$$

$$VOP = lgPTS / (lgFGA - lgORB + lgTO + 0.44 \times lgFTA). \tag{3}$$

$$DRBP = (lgTRB - lgORB) / lgTRB. \tag{4}$$

Once uPER is calculated, it must be adjusted for team pace and normalized to the league to become PER:

$$PER = uPER \times (lgPace/tmPace) \times (15/lguPER). \tag{5}$$

The meanings of symbols are defined in Table 1.

4 Optimized PER Formula

4.1 The Formula

Based on the weighted integration of the players' technical statistics, the official PER of NBA balances the variation of different positions and different time and eliminates

the influences caused by different strategies and teams' general situation. To a certain extent, it correctly reflects the varied technical demands in the fields. However, due to its excessive focusing on the fairness and consistency, it fails to present the difference between each position effectively. For this reason, this paper evaluates the correlation between every factor and PER on the basis of the official computing formula of PER, and revises the original formula by weighting with correlation coefficients to get the new optimized efficiency rating formula for five positions—PF, SF, C, PG and SG.

Table 1. Symbols used in the PER formula

Symbols	Meanings	Symbols	Meanings	Symbols	Meanings
AST	assist	STL	steal	TO	turnover
BLK	block shot	TRB	total rebound	ORB	offensive rebound
FG	field goal	PF	personal foul	FT	free throw
FGA	field goal attempts	3P	three-point goal	FTA	free throw attempts
MP	players playing time	lgTO	conference average turnover	lgPTS	conference average points
DRBP	conference average offensive rebound	VOP	conference average offensive efficiency	lgTRB	conference average total rebound
tmAST/ tmFG	team assist / team field goal	lgFGA	conference average field goal attempts	lgORB	conference average offensive rebound
lgFTA	conference average free throw attempts	lgFT/ lgPF	conference average free throw / conference average foul	lgAST/lgFG	conference assist / conference field goal
lgFTA/ lgPF	conference average free throw attempts / conference average personal foul	lgFG/ lgFT	conference average field goal/conference average free throw	lgPace/ tmPace	conference average number of rounds /team average number of rounds

The formula of unadjusted PER (uPER) is revised into:

$$uPER=(1/MP)\times\{\omega_1\times 3P+\omega_2\times(2/3)\times AST + (2-\omega_3\times factor\times tmAST/tmFG) \times\omega_4\times FG + \omega_5\times FT\times 0.5\times[1+(1-\omega_3\times tmAST/tmFG)] + \omega_3\times(2/3)\times tmAST/tmFG - \omega_6\times VOP\times TO - \omega_7\times VOP\times DRBP\times(FGA-FG) - \omega_8 \times VOP\times 0.44\times[0.44+ (0.56\times DRBP)] \times(FTA - FT) + \omega_9\times VOP\times(1-DRBP)\times(TRB-ORB) + \omega_{10}\times VOP\times DRBP\times ORB +\omega_{11}\times VOP\times STL + \omega_{12}\times VOP\times DRBP\times BLK - \omega_{13}\times(PF\times lgFT/lgPF) - 0.44 \times lgFTA /lgPF \times VOP \}. \quad (6)$$

Table 2 shows the correlation coefficients between different factors and PER values computed by SPSS.

Table 2. The correlation coefficients of different positions between each factor and the PER values

correlation coefficients	PF	SF	C	PG	SG
ω_1	0.198	0.173	0.185	0.130	0.309
ω_2	0.650	0.579	0.658	0.679	0.542
ω_3	0.068	0.036	-0.056	0.135	0.098
ω_4	0.799	0.760	0.886	0.802	0.676
ω_5	0.759	0.735	0.813	0.838	0.723
ω_6	0.608	0.668	0.736	0.690	0.651
ω_7	0.680	0.600	0.766	0.663	0.598
ω_8	0.623	0.675	0.622	0.646	0.583
ω_9	0.786	0.565	0.792	0.629	0.389
ω_{10}	0.607	0.483	0.693	0.546	0.517
ω_{11}	0.505	0.413	0.544	0.578	0.553
ω_{12}	0.423	0.446	0.455	0.257	0.220
ω_{13}	0.358	0.347	0.412	0.360	0.344

The optimized PER formula (6) uses the above correlation coefficients. Furthermore, all the PERs of the 306 players of NBA can be computed with model (6) in the 2010-2011 seasons. Table 3 shows some typical players' PER.

Table 3. PER value of some typical NBA players before and after optimization

Name	PER (before)	PER (after)	Name	PER (before)	PER (after)	Name	PER (before)	PER (after)
Luis Scola	18.68	18.63	Chris Paul	23.92	21.48	Deron Williams	21.69	20.78
David Lee	17.14	17.73	Al Jefferson	19.65	20.59	Tim Duncan	21.37	21.66
Roy Hibbert	16.14	16.96	Steve Nash	22.36	19.63	Derrick Rose	22.78	20.85
Nenê Hilario	20.53	21.45	Kobe Bryant	24.00	23.78	LeBron James	26.86	23.13
Andrew Bynum	20.81	21.80	Blake Griffin	22.51	22.11	Dirk Nowitzki	23.75	22.89
Dwight Howard	26.40	27.59	Kevin Durant	23.95	21.16	Dwyane Wade	25.04	23.72
Carmelo Anthony	21.23	22.77	Pau Gasol	23.71	22.61	Danny Granger	17.86	15.74
Amar'e Stoudemire	23.77	25.22	Russel Westbrook	23.91	19.50	LaMarcus Aldridge	21.82	21.33

Observed from Table 3, it can be seen that NBA players' PERs ameliorate to different extent. Take Amar'e Stoudemire as example. In this season he joined a new team and dramatically improved the performance of his new team by his stable and high-effective skills. Therefore, his PER value enhances after optimization for his great contribution to the new team. Thus, the official PER value of Amar'e Stoudemire is underestimated as Table 3 shows. While for LeBron James and Dwyane Wade, as part of the "Big Three" of Miami Heat, they win more games for their teams and bring themselves the "perfect" technical statistics by their intimate cooperation. What's more, LeBron James himself led Cleveland Cavaliers to acquire the best achievements for the last several seasons. However, the "Big Three" of Miami Heat only achieve the second place in the Eastern association in this season. That is to say, compared with Amar'e Stoudemire, their contributions to team is a little lower, thus the optimized PER values decrease a little. Thus, the official PER values are overestimated by the official as Table 3 shows. From the above analysis, the optimized computing model of PER is more pertinent and could reflect the players performances and contributions to their teams more accurately.

4.2 Comparison of Optimized Results

The purpose of the optimizations model is to underline the different technical demands for the different positions of the match in the PER formula. Therefore, this paper adopts the method of weight comparison to compare each factors in original and the revised PER formula, in order to determine whether factors which have high relationship with players PER have been improved and to evaluate the results of optimization under this criterion. Since PER and uPER relate to each other directly, the comparison of factors in uPER formula could also bring effective results. At first, we simplify the uPER formula as below:

$$uPER=c_1\times 3P+c_2\times AST+c_3\times FG+c_4\times FT+c_5\times tmAST/tmFG+c_6\times TO+c_7\times (FGA-FG)+c_8\times(FTA-FT)+c_9\times(TRB-ORB)+c_{10}\times ORB+c_{11}\times STL+ c_{12}\times BLK+c_{13}\times PF+c_0 \,. \tag{7}$$

During simplifying, factors VOP, DRBP, lgFT/lgPF and lgFTA/lgPF could be known by computing the conference statistics in this season, whose values are shown in Table 4.

Table 4. The results of some factors in uPER formula

factor	VOP	DRBP	lgFT/lgPF	lgFTA/lgPF
0.408861	0.765227	0.358842	0.965308	1.053977

Then, we normalize those weight factors by the formula below:

$$\mathbf{c_i}=|\,\mathbf{c_i}\,|/\textstyle\sum|\,\mathbf{c_i}\,| \tag{8}$$

Because of the relationship between some weight factors in the uPER and the general technical statistics of teams, different players' weight in uPER formula are not

identical (the players in the same team have the same weight). Therefore, one representative player from five positions is chosen respectively, as shown in Figure 1 to Figure 5.

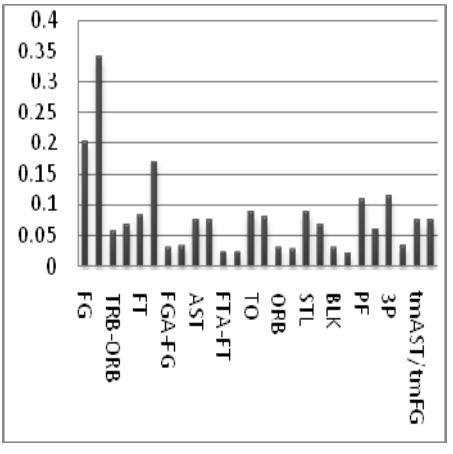


Fig. 1. Amar'e Stoudemire (PF)

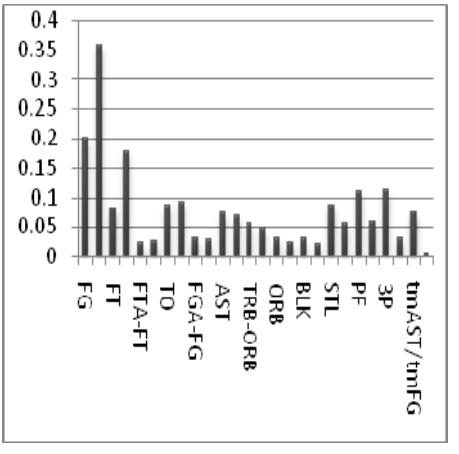


Fig. 2. Kevin Durant (SF)

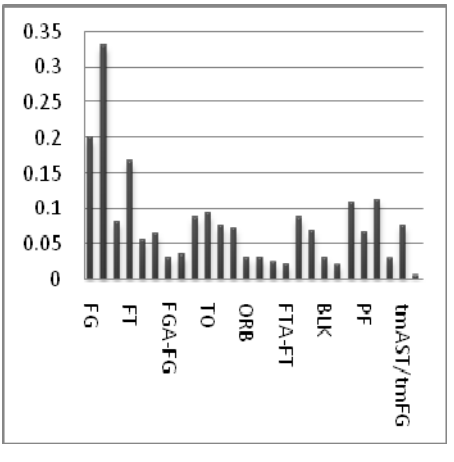


Fig. 3. Dwight Howard (C)

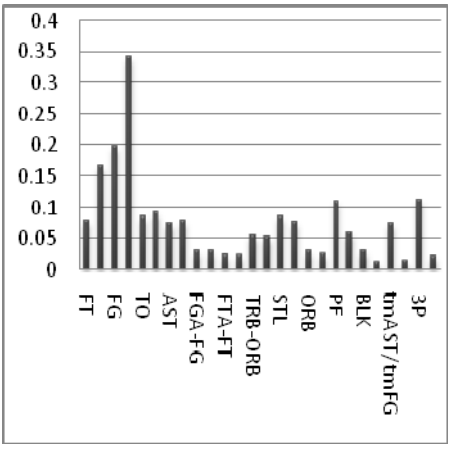


Fig. 4. Chris Paul (PG)

These five above figures show the values contrasted between the original ones and weight-revised ones of 13 factors in the PER formula of different positions. Take Kobe Bryant as example to analyze. The first column represents the weight value of free throw before optimization. The second column is the weight value of free throw after optimization. The third column represents the weight value of field goal before

optimization. The forth column is the weight value of field goal after optimization. It can be seen that the weights of the free throw and the field goal should be pay more consideration for a shooting guard.

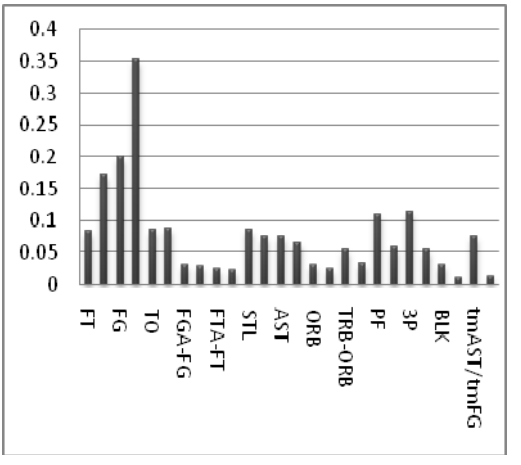


Fig. 5. Kobe Bryant (SG)

In previous discussion, the weight values of each position after and before optimization have been compared. Then they will be compared among different positions in Table 5, which shows the values of the normalized weight factors of each position. Then the results of Table 5 will be analyzed when different positions' responsibilities in the basketball match are concerned.

Table 5. The value of weight factors of each position

Factor	PF	SF	C	PG	SG
3P	0.033966	0.031097	0.0302550	0.022531	0.054927
AST	0.074337	0.069384	0.0717390	0.078453	0.064230
FG	0.340449	0.356810	0.3291486	0.341079	0.351782
FT	0.168312	0.178008	0.1660680	0.166529	0.173193
tmAST/tmFG	0.007777	0.004314	0.0061060	0.015598	0.011613
TO	0.079814	0.091884	0.0921070	0.091510	0.088552
FGA-FG	0.032032	0.029615	0.0343992	0.031553	0.029189
FTA-FT	0.023064	0.026185	0.0219530	0.024162	0.022365
TRB-ORB	0.066155	0.049829	0.0635488	0.053486	0.033926
ORB	0.028593	0.023840	0.0311210	0.025985	0.025235
STL	0.066293	0.056809	0.0680790	0.076656	0.075222
BLK	0.019926	0.022014	0.0204330	0.012231	0.010738
PF	0.059283	0.060210	0.0650410	0.060228	0.059027

Point guard is the person who gets the ball in most chances. He must be able to steal the ball, bring it across the court under defense and then pass it to the place where the ball should be. To be specifically, he needs to organize the team's attack, facilitating his teammates' attacking. Just as shown in the table above, his weight of steal and assist should be the highest of the 5 positions.

Shooting guard's main mission is to gain points. He is the second point-winner only after Small Forward[10]. Shooting guard needs good chance to shoot and so he should have excellent accuracy and stability of shooting.

Small forward is the most important point-winner[10], thus having a higher value on FG.

Power forward's responsibilities are: backboard and defense. Therefore, the data relating to backboard of power forward is higher than other positions.

As the name indicates, center is the central person of a team. He is crucial in both attack and defense. So his statistics of backboard and block shot is relatively high.

The factor of tmAST/tmFG is related to the team situation. Although the integral strategy of a team influences the players' performances to a certain extent, the personal data should be more emphasized while evaluating players' personal Efficiency Rating. Thus this factor is a little lower for all positions.

From the above analysis, weight values in optimized formula of each position match well with functions and technical demands of each position in basketball match. Therefore, the optimized model is more targeted and practical than the NBA's official formula of player efficiency rating computation.

5 Conclusion and Suggestions

1. Optimized PER formula effectively indicates the technical demands of different positions, which is of certain guidance in computing players' PER in the future.
2. According to the relationship and weights of technical factors in PER formula, NBA players could have some special training to improve their PER. NBA teams should also make some technical training programs suitable for players in order to improve their performances and more wins.

Acknowledgements

This paper is supported by the Innovation Lab of Mathematics and Physics of Beijing University of Posts and Telecommunications and the Fundamental Research Funds for the Central Universities (BUPT2009RC0701) of China.

References

1. Xue, L.: Discussion on the Position Technique in Basketball Competition. J. China Sport Science and Technology 38(7), 31–33 (2002)
2. Calculating PER, http://www.basketball-reference.com/about/per.html
3. http://www.hoopchina.com/stats/

4. http://china.nba.com/
5. http://nba.espnstar.com.cn/
6. Xi, J., Zhu, W.: Excel 2007 From Novice to Professional. Tsinghua University Press, Beijing (2008)
7. Science and Education Workshop. Excel 2007 Functions, Formulas and Charts.Tsinghua University Press, Beijing (2008)
8. Ding, G., Li, T.: SPSS Statistics Tutorial - from research design to data analysis. Machinery Industry Press, Beijing (2005)
9. Lu, W.: SPSS statistical analysis. Electronic Industry Press, Beijing (2010)
10. Introduce the main location of the basketball court, http://wenku.baidu.com/view/4277602ded630b1c59eeb598.html

A Reliable Classification Method for Paper Currency Based on LVQ Neural Network

Hongling Gou, Xiaofeng Li, Xuedong Li, and Jing Yi

School of Information Science and Technology,
Shandong Jinzhou University, Jinan, 250101
{gouhl,sdlixiaofeng,lixuedong,yj_yijing}@sdjzu.edu.cn

Abstract. To increase the reliability of currency classification, a classification method using neural networks with multi-pattern vectors is proposed in this paper. The data space of samples are divided into three blocks, then the latter are further divided into four sub-pattern vectors, and kernel principal component analysis is applied to extract features and assemble feature vectors to train LVQ neural network classifier. We draw the conclusion by testing new fifth edition RMB including four kinds of inputting directions of 1 Yuan, 5 Yuan, 10 Yuan and 20 Yuan RMB, up to 800 samples that PCA can compress data and decrease dimension of input vectors, extract the feature vectors effectively, thus the high-level reliability can be achieved by using the LVQ network classifier.

Keywords: Principal Component Analysis, sub-pattern vector, neural network classifier.

1 Introduction

Artificial neural network has been widely used in pattern recognition, and the currency feature extraction has direct influence on the scale and classifying performance of neural network, the former of which is considered as one of the most focus-worthy questions in pattern classification [1, 2]. To solve the scale problem of neural network, the mask technology[2] and the method of principal component analysis[3,4] had been proposed to apply to Extract the characteristics of paper currency by the previous researchers, Takeda, F and Ahmadi, A.

In this paper, the currency pattern vectors are divided into different blocks, and which is partitioned into four regions namely sub-patterns, and principal component analysis is implemented on each sub-pattern, and the feature vectors of each sub-pattern can be acquired by projection on each principal component, then they are combined and inputted into LVQ neural network classifier. By testing new fifth edition RMB of 1 Yuan, 5 Yuan, 10 Yuan, 20 Yuan including four kinds of inputting directions, up to 800 samples, we can draw the conclusion that the classifying performance is reliable by combining multi-pattern feature vectors and applying LVQ neural network classifier.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 243–247, 2011.

2 Data Processions

Six optical sensors are applied which are laid beside the banknotes in bilateral symmetry, and red light, infrared light, ultraviolet light are used in turn to irradiate the surface of the banknotes. The corresponding reflection and refraction are received and recorded and 36 sets of data are generated. Because banknotes with different currency values have different lengths, the numbers of sample points are different which are 36, 38, 40 and 42. To calculate the eigenva1ues and eigenvectors of such large-scaled matrices, compute cost will be increased, and the speed of currency recognition will slow down inevitably.

After extensive tests, 12 sets of data which best represent the features of the banknotes are selected, and each set is quantized with the interval of 0x0~0xFF, each of which corresponds a wave curve.

3 Pattern Vector Partition and Feature Extraction

3.1 Pattern Vector Partition

In this paper, 12 sets of data from each currency value banknote are selected for feature extraction, and they are further divided into three data blocks, the red light data block, the infrared light data block and the ultraviolet light data block which correspond to the pattern vectors of X_1, X_2, X_3. The dimension of an unknown sample pattern vector cannot be determined, whilst the same dimension of inputted pattern vectors is required in data processing.

In this paper, the pattern vectors of X_1, X_2, X_3. are divided into four sub-pattern vectors also $\boldsymbol{X}_k = [\boldsymbol{X}_{k1}, \boldsymbol{X}_{k2}, \boldsymbol{X}_{k3}, \boldsymbol{X}_{k4}]^T$, $\boldsymbol{X}_{ki} = \left[\boldsymbol{x}_{(i-1)*36}, ..., \boldsymbol{x}_{i*36} \right]^T$, $k = 1,2,3, i = 1,2,3,4$, the dimension of sub-pattern is determined based on the dimension of the pattern vector of 1 Yuan which is 36, the first 36 data in each sub-pattern are reserved, thus dimension of each sub-pattern is 36, and the dimension of X_1, X_2, X_3. is 144. Thus different sample data spaces with different currency values are divided into three pattern vectors with the same dimension, and each pattern vector is re-divided into 4 sub-pattern vectors with the same dimension.

3.2 Feature Extraction Based on PCA Algorithm

The larger the variance of a random variable is, the more information it carries. Principal component analysis in Gaussian distribution equals to maximizing the amount of information in output signal, and the number of principal component is determined according to the variance contribution rate of each component. The variance contribution rate of the *i*th principal component is defined as $\eta_i = \lambda_i / \sum_{k=1}^{p} \lambda_k$, and the contribution rate of accumulative variance of the first m principal components is $\eta = \sum_{i=1}^{m} \lambda_i / \sum_{i=1}^{p} \lambda_i$.

In PCA algorithm, the eigenvalues of covariance matrix C in data space are ordered as follows, $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p)$, and the cumulative variance contribution of the first *m* principal components $\eta \geq 0.80$ in this paper, and by principal component analysis the input vector of *p* dimension can be transformed into the principal component of *m* dimension. The minimization of errors can be achieved by restructuring the original data with the principal component of *m* dimension. Here λ_i plays as the eigenvalue of covariance matrix C of random vector *X*.

According to the above definition, the cumulative variance contribution rates of each sub-pattern vector are computed. After analysis, the cumulative variance contribution of the first 5 principal components in each sub-pattern are all greater than 0.80. The PCA algorithm is implemented on the 12 sub-pattern vectors, as a result, there are totally 46 feature data extracted as currency feature vectors inputted into network.

4 Currency Classifications

4.1 Learning Vector Quantization Algorithm

Learning vector quantization or LVQ has been wildly used in neural network pattern recognition due to its own advantages such as automatically adapting to learning and storing message, simple network structure, proceeding high-dimension input with lower computing cost.

Let X^p represent the *p*th training pattern vector, and T^p represent the class to which X^p belongs, and C_i is the class which is represented by the *i*th output neuron, and the neuron number of the hidden layer is *n,* and the procedure of the learning vector quantization algorithm is as follows.

Step 1, initialize the weight vector, $W = \{w_1, w_2, \cdots, w_n\}$, and initialize the learning rate $\alpha \in [0,1]$

Step 2, compute the distances from each sample vector X^p to the neurons of the competition layer and the winner neuron is selected as the neuron with the minimum distance.

$$\left\| w_k - X^p \right\| < \left\| w_i - X^p \right\| \quad \text{and} \quad i = 1,2,\cdots,n \tag{1}$$

e weight is modified as follows,

$$w_k(t+1) = \begin{cases} w_k(t) + \alpha(t)\left[X^p - w_k(t)\right] & \text{if} \quad T^p = C_k \\ w_k(t) - \alpha(t)\left[X^p - w_k(t)\right] & \text{if} \quad T^p \neq C_k \\ w_k(t) & \text{others} \end{cases} \tag{2}$$

Step 3, modify the learning rate $\alpha(t)$, $\alpha(t)$ is a iterative decrease function.

$$\alpha(t) = \alpha_0 \times \left(1 - \frac{t}{T}\right) \tag{3}$$

In the formula, α_0 is the initialized learning rate, and $\alpha_0 \in [0,1]$, and T is training iteration time. The formula means the iteration learning rate in the t time gradually decreases as the training times increase.

Step 4, examine the termination condition, and withdraw if certain conditions prevail, otherwise go back to step 2 to continue the procedure.

4.2 The Result of Currency Classification

The numbers of input and output neurons are 46 and 16 by its feature vectors and pattern class, and the number of nodes in the hidden layer is 3 to 6 times as large as the one in the output layer. In this paper, the number of nodes in the hidden layer is 3,4,5,6 times as large as the one in the output layer, namely, the numbers of neurons are 48, 64, 80, 96, and $\alpha_0 = 0.001, T = 400$. ten samples are selected from each currency value of 1 Yuan, 5 Yuan, 10 Yuan, 20 Yuan, 40 samples with four directions A, B, C and D compose a training set with 160 column vectors to train LVQ neural network classifier, and the target vectors are the 160 column vectors with 16 dimensions; the testing set consists of 800 samples of each currency value RMB with four directions A, B, C and D. Recognition rate, rejection rate and false acceptation rate which serve as evaluation indexes of neural network classifier performance are shown as follows, *R1*, *R2*, *R3* are recognition rate, rejection rate and false acceptation rate.

Table 1. The classification results of LVQ neural network

	LVQ1(46×48×16)			LVQ2(46×64×16)		
¥	*R1*	*R2*	*R3*	*R1*	*R2*	*R3*
1	98	0	2	98	0	2
5	100	0	0	100	0	0
10	100	0	0	100	0	0
20	100	0	0	100	0	0

From the table 1, the overall recognition results of network LVQ1 and LVQ2 keep the same. There are 4 false recognitions on the B side of 1 Yuan RMB in the two networks because each pattern vector category learns with the same weight in training process, namely, the relevant 3 and 4 code vectors.

Table 2. The classification results of LVQ neural network

	LVQ3(46×80×16)			LVQ4(46×100×16)		
¥	*R1*	*R2*	*R3*	*R3*	*R3*	*R3*
1	98.5	0	0	100	0	0
5	100	0	0	100	0	0
10	100	0	0	100	0	0
20	100	0	0	100	0	0

From the table 2, In network LVQ3, the weight of the B side of 1 Yuan RMB is increased and the code vectors of 6, 14, 6, 6 are assigned to the A, B, C, D pattern category, while RMB with other currency values are assigned 4 code vectors. The result is the false acceptation rate of the B side of 1 Yuan RMB in LVQ3 decreased to 0, while the one of the C side is 10%. In LVQ4, 5, 16, 14, 5 code vectors are assigned to the A, B, C, D pattern category of 1 Yuan RMB, and the other pattern categories are assigned 5 code vectors respectively, the false recognition rate of LVQ4 decreases significantly to 0, and the recognition rate is 100%.

5 Conclusions

As the main limitation of PCA is its global linearity, that is, it only defines a linear projection of data and does not model nonlinear relationship among variables [5] and it depend largely on linear projection and second-order statistics, thus we propose the method of dividing pattern vectors into sub-pattern vectors to minimize the limitation as well as PCA algorithm. PCA algorithm can be implemented on each sub-pattern vectors, and feature vectors can be combined and inputted into LVQ network classifier. The experiments have indicated that PCA algorithm implemented on each sub-pattern vectors can effectively compress data and extract currency character of feature vectors, and the overall performance can be improved by assigning different code vectors to different categories.

References

1. Teranishi, M., Omato, S., Kosaka, T.: New and Used Bills Classification For Cepstrum Patterns. In: International Joint Conference on Neural Networks, IJCNN Apos, vol. 6, pp. 3978–3981 (July 1999)
2. Takeda, F., Nishikage, T., Matsumoto, Y.: Characteristics Extraction of Paper Currency Symmetrical Masks Optimized by GA and Neuro-Recognition of Multi-National Paper Currency. In: IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on Neural Networks Proceedings, May 4-8, vol. 1, pp. 634–639 (1998)
3. Ahmadi, A., Omatu, S., Kosaka, T.: Improvement of the Reliability of Bank Note Classifier Machines. In: 2004 Proceedings of the IEEE International Joint Conference on Neural Networks, July 25-29, vol. 2, pp. 1313–1316 (2004)
4. Zhang, B., Fu, M.: A nonlinear neural network model of mixture of local principal component analysis: application to handwritten digits recognition. Pattern Recognition 34(2), 203–214 (2001)
5. Ahmadi, A., Omatu, S., Kosaka, T.: A Methodology to Evaluate and Improve Reliability in Paper Currency Neuro-Classifiers. In: Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation, 2003, July 16-20, vol. 3, pp. 1186–1189 (2003)

Wavelet Analysis in the Application of Rotor Flowmeter Reading

Hong-bo Zhang and Yu-ming Shen

School of Optical-Electrical and Computer Engineering,
University of Shanghai for Science and Technology,
Shanghai 200093, China
lilygirl2008@sina.com

Abstract. Different from traditional spatial or frequency domain image denoising and image enhancement method, discussed the characteristics of rotor flowmeter image, then tried to use the wavelet denoising and wavelet enhancement for its pretreatment, and further realized the rotor flowmeter reading. The wavelet-denoising and wavelet-enhancement takes the wavelet decomposition and restructuring as the foundation, then the effect of wavelet decomposition and restructuring is closely related with the choice of wavelet base. Therefore, comprehensively compared the commonly used wavelet base both from the theory and experiment view, and indicated the wavelet base which most matches with the above application.

Keywords: Wavelet base; Wavelet-denoising; Wavelet-enhancement; Rotor flowmeter.

1 Introduction

The realization of the rotor flowmeter reading is divided into three stages orderly: image acquisition, image processing, rotor corresponding scale recognition. The rotor corresponding scale identification adopted the peak method. Because of the influence of specific environment which rotor flowmeter be applied in, the obtained image's quality is uneven, containing noise (such as the white Gaussian noise, random white noise). Therefore, before identifying the rotor corresponding scale, image processing is required. Image processing mainly includes denoising and enhancement, usually in the spatial or frequency domain [1] [2]. According to the characteristics of rotor flowmeter image and advantages of wavelet multi-resolution analysis, this paper attempted to use the wavelet-based denoising, enhancement. The actual operation of the above two approaches makes wavelet decomposition and reconstruction as the foundation, and the quality of wavelet decomposition and reconstruction is closely related to the choice of wavelet base. Therefore, this paper focused on this particular image signal of rotor flowmeter, decided which wavelet base will be the best to obtain accurate wavelet decomposition, reconstruction, and made necessary experimental comparison by the Matlab wavelet toolbox and other.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 248–253, 2011.

2 The Basis of Wavelet Base Choosing

From universal sense, wavelet base has the following five key indicators [3][4] : symmetry, orthogonality (linear phase characteristic), compact support, vanishing moments, regularity. However, apart from Haar wavelet, the symmetric orthogonal wavelet with linear phase does not exist, so generally relaxed conditions to use biorthogonal wavelet. Image processing generally selects biorthogonal wavelet bases, such as biorNr.Nd, rbioNr.Nd [5].

From the above analysis and the importance of accurate wavelet decomposition and reconstruction for this application, the wavelet base, which suitable for the application, must have symmetry, biorthogonality, good regularity, compact support, and the best one is the biorthogonal wavelet base.

In this application of rotor flowmeter reading, selecting the wavelet base, in addition to considering the above universal index, but also with specific features of the rotor flowmeter image and the precision of final reading to comprehensively assess. Therefore, this paper customized three evaluation index:

(1) PSNR (peak signal-to-noise ratio) of wavelet denoised image

Wavelet denoising through the low frequency part reconstruction after decomposition of original image to realize, denoising effect can be measured by the reconstructed image's PSNR.PSNR represents the difference between two images, is one of the measure standard of the quality of image reconstruction, but also an important indicator of evaluating image processing (such as denoising) method .With the premise that denoising using wavelet reconstruction in all cases, to experiment with a variety of wavelet bases, the value of PSNR of final denoised image can be used as the evaluation criteria of wavelet base stand or fall. The greater the value of PSNR is, the better the quality of reconstructed denoising image is, the corresponding wavelet is more suitable for this application.

For 256-color grayscale image, peak signal to noise ratio is calculated [4] [6] as follows:

$$PSNR = 10 \log 10 \frac{255^2}{D}. \tag{1-1}$$

$$D = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [x(i, j) - \hat{x}(i, j)]^2 . \tag{1-2}$$

Type (1-2), the M, N respectively is the number of image elements in each row and column, $x(i, j)$ is the gray value for the original image at point (i, j), $\hat{x}(i, j)$ is the gray value for the reconstructed image at that point .

(2) The feasibility of rotor corresponding scale identification

Wavelet analysis in the application of rotor flow meter reading, after wavelet decomposition, wavelet denoising, wavelet enhancement and edge detection, with the help of peak method[7] [8] to identify rotor corresponding scale line. First of all,

introducing the peak method in detail, refer to Figure 1: 1 to 530 of abscissa corresponding with 1 to 530 row of original image (the original image's size is 530 rows 70 columns), y-coordinate value is the accumulate and of corresponding row pixel values in the image after edge detection. Known that rotor is constitutes by nearly cylindrical upper part and tapered bottom part, and the scale of rotor largest cross-section corresponding to equals the rotor flowmeter reading, also equals the abscissa which corresponded with the maximum y-coordinate in peak method waveform figure . All of the following rotor corresponding scale measured values are obtained based on this method. Therefore, the good peak method waveform figure should have the maximum y-coordinate in where the abscissa corresponding to rotor largest cross-section.

In the premise that all other conditions must be the same, because the selected wavelet base is different, the final obtained peak method waveform figure is different, good peak method waveform figure corresponding to good wavelet base (only for this application). Only with a good peak method waveform figure, the identification of rotor corresponding scale line is feasible.

(3) Calculation speed

Considering the real-time requirement of actual application, especially made calculation speed as one of indicators to evaluate wavelet base stand or fall [9]. Specifically, the calculation speed measured by the amount of time spent, the shorter the time spent is, the better the wavelet base is.

3 Experimental Comparison

Experimental platform, mainly including MATLAB and its image processing toolbox and wavelet toolbox, LZB 6 rotor flowmeter (measuring range 0.06-0.6, minimum scale 0.01, units are $^{Nm}/_{n}$), a industrial camera. Accessed rotor flowmeter image is 256-color BMP grayscale image, the size is 530 rows 70 columns, see Figure 2.

Using Matlab wavelet toolbox function wavedec2() to realize the multi-resolution two-dimensional discrete wavelet decomposition, this function detailed introduction as follows: [C, S] = wavedec2 (X, N, 'wname'); input parameters X, N, 'wname' respectively is the original image which will be decomposed, decomposition level, the name of used wavelet base. Output parameters C, S separately is decomposed coefficient matrix, marking matrix, the former can fully express the all characteristic of original image X. In this paper, the X is the rotor flowmeter image, which is known certainly. The experiment shows that the general decomposition level N = 2.

Here on the variable parameter 'wname' with different settings: (a) wname = bior1.1;(b) wname = bior6.8;(c) wname = coif1;(d) wname = coif4;(e) wname = db15;(f) wname = db4;(g) wname = dmey;(h) wname = haar;(i) wname = rbio1.1;(j) wname = rbio6.8;(k) wname = sym2;(l) wname = sym10 and through the above custom criteria to compare each other.

(1) PSNR of wavelet denoised image

Known M = 530, N = 70. The gray value $x(i, j)$ of original rotor flowmeter image at point (i, j) and gray value $\hat{x}(i, j)$ of denoised rotor flowmeter image at that point, also

is known. Referenced type (1-1), (1-2) and made calculation, the PSNR value corresponding to selected wavelet base can be obtained, the result showed in Table 1 .

(2) Calculation speed

Calculation speed, measures by the spent time of relevant processing. In order to improve the accuracy and comparability, the spent time is the average time of nine times repetitive measuring, the results showed in Table 1.

(3) The feasibility of rotor corresponding scale identification

Choosing different wavelet base got different peak method waveform figure, shown in Figure 3. Characteristics of peak method waveform figure (in Fig. 3) and comparsion of each other, see Table 1.

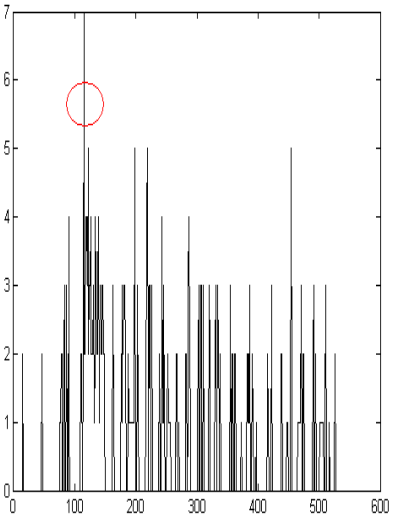


Fig. 1. Example of peak method waveform figure

Fig. 2. To be processed rotor flowmeter image

Observing be processed rotor flowmeter image (Fig. 2), rotor corresponding scale line can be visually marked closely to the line but not exceeding 120 in the image. From Table 1, considering the abscissa of maximum value in the peak method waveform figure, and only (b) bior6.8, (j) rbio6.8, (l) sym10 can meet the requirements. What's more,(b) bior6.8, (j) rbio6.8, (l) sym10: (j) rbio6.8 's spent time 0.7151 is the shortest, slightly smaller than (b) bior6.8 's 0.7246; from the PSNR perspective, it is only slight difference between the two, are higher than (l) sym10.Therefore, considering the three custom indicators, the wavelet bases which suitable for this application best are (b) bior6.8 and (j) rbio6.8.

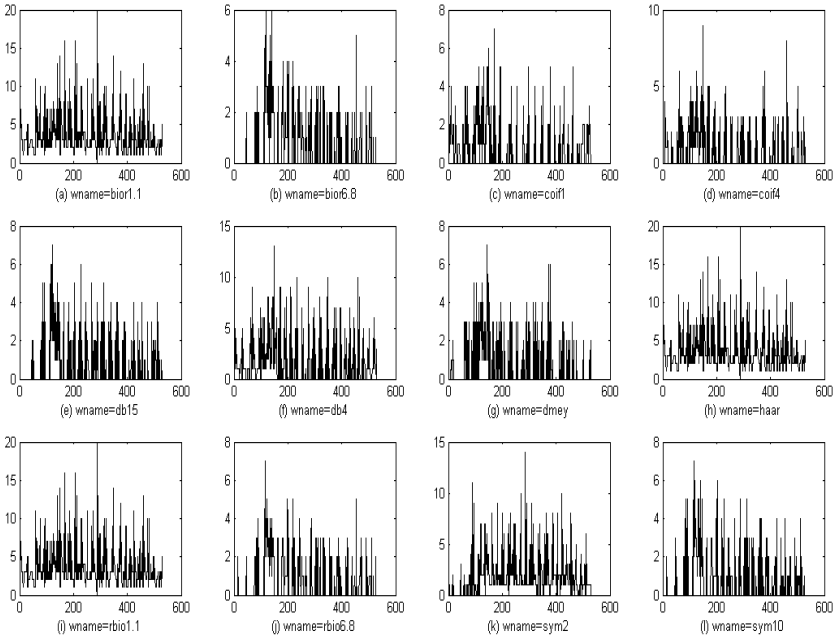


Fig. 3. Peak method waveform figure

Table 1. Analysis of peak method waveform figure, the spent time, PSNR

parameters / wavelet base	peak method waveform figure		spent time (Units: seconds)	PSNR
	maximum value	abscissa of maximum value		
(a) bior1.1	exist	288	0.6540	62.5067
(b) bior6.8	exist	119	0.7246	63.1027
(c) coif1	exist	170	0.5983	62.7979
(d) coif4	exist	149	0.8454	62.9240
(e) db15	exist	123	1.0205	63.0530
(f) db4	exist	149	0.5855	62.8434
(g) dmey	exist	144	4.6865	62.9836
(h) haar	exist	288	0.5120	62.5067
(i) rbio1.1	exist	288	0.5277	62.5067
(j) rbio6.8	exist	115	0.7151	63.1003
(k) sym2	exist	286	0.5488	62.9019
(l) sym10	exist	115	0.8841	63.0536

4 Summary

This paper adopted a new based on wavelet analysis method to finish the rotor flowmeter reading, for this particular image signal of rotor flowmeter, focused on choosing the best wavelet base to acquire accurate wavelet decomposition and reconstruction, did detailed comparison by the MATLAB simulation experiment and analyzed the results. Owing to space restrictions, the wavelet bases used in the paper just a part of so much wavelet, and the comprehensive of experiment will be strengthened. In view of the comparability of principle, the method of this paper can be used for spherical rotor flowmeter readings 、 level readings, etc.

References

1. Gonzalez, R.C., Yuan, Q.-Q., et al.: Digital Image Processing, 2nd edn. Electronic Industry Press, Beijing (2007)
2. Gonzalez, R.C., Yuan, Q.-Q., et al.: Digital Image Processing (MATLAB version). Electronic Industry Press, Beijing (2005)
3. Qin, D.-Y., Wang, J.-D., Li, P.: Selection and Analysis of Wavelet Base in the Image Fusion. Photoelectron Technology 26(3), 203–207 (2006)
4. Ke, L., Huang, L.-Q.: Selection of Wavelet Base for the Remote Sensing Image Real-Time Compression. Optical Technology 31(1), 77–80 (2005)
5. Guo, Y.: Selection of Wavelet Base for Vibration Signal Processing. University of Hefei for Technology, Hefei (2003)
6. Cai, D.-H., Yi, X.-M.: Influence of Wavelet Base Selection to Denoising of the Image. Math. 25(2), 185–190 (2005)
7. Cao, H.: Research of Device for Level Automatic Reading Based on Image Processing. University of Shanghai for Science and Technology, Shanghai (2009)
8. Cao, X.-X., Shen, Y.-M.: Recognition Of Flow Reading For Float Flowmeters Based On Digital Image Technique. In: The 15th International Flow Measurement Conference, FLOMEKO 2010, p. 118 (2010)
9. Yan, Y.: Electrical Fault Diagnosis and the Choice of Wavelet Base. University of Taiyuan for Science and Technology, Taiyuan (2006)

Prediction of Protein-Protein Interactions Using Local Description of Amino Acid Sequence

Yu Zhen Zhou[1], Yun Gao[1], and Ying Ying Zheng[2]

[1] Computer Engineering Dept., JiangSu College of Information Technology, Wuxi 214181, China
[2] Institute of Intelligent Machines, Chinese Academy of Sciences,Hefei 230031, China
zyz9508@sina.com, xiaoxiao@000416@yahoo.com.cn, yingzheng1982@hotmail.com

Abstract. Protein-protein interactions (PPIs) are essential to most biological processes. Although high-throughput technologies have generated a large amount of PPI data for a variety of organisms, the interactome is still far from complete. So many computational methods based on machine learning have already been widely used in the prediction of PPIs. However, a major drawback of most existing methods is that they need the prior information of the protein pairs such as protein homology information. In this paper, we present an approach for PPI prediction using only the information of protein sequence. This approach is developed by combing a novel representation of local protein sequence descriptors and support vector machine (SVM). Local descriptors account for the interactions between sequentially distant but spatially close amino acid residues, so this method can adequately capture multiple overlapping continuous and discontinuous binding patterns within a protein sequence.

Keywords: Protein-protein interactions; Protein sequence; Local descriptors; SVM.

1 Introduction

Protein-protein interactions (PPIs) play important roles in most cellular processes, such as transcription regulation, signal transduction [1], and recognition of foreign molecules. Knowledge of PPIs can provide insight into protein functions [2, 3], lead to a better understanding of disease mechanisms and suggest novel methods for designing drugs that modulate specific disease pathways. In recent years, high throughput technologies have been developed for the large-scale PPI analysis, such as yeast two-hybrid screening methods [4], immunoprecipitation [5], and protein chips [6]. However, there are some disadvantages of existing experimental methods, such as time-intensive, high cost and a small fraction of the complete PPI network covered. In addition, these approaches suffer from high rates of both false negative and false positive predictions. Therefore, there is a strong motivation to develop reliable computational methods for inferring protein interactions [7], which provide an attracting perspective on predicting and understanding PPIs as complementary methods to experimental ones.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 254–262, 2011.

A number of computational methods [7] have been developed for the prediction of PPIs based on various data types, including genomic information, protein domain and protein structure information. However, these methods are not universal, because the accuracy and reliability of these methods depend on the prior information of the protein pairs such as the information of protein homology [8, 9]. Moreover, compared to the rapid increase of the number of protein sequences, the protein three-dimensional structure data is scarce. So approaches that derive information directly from amino acid sequence information are of particular interest [8-14]. Many groups have engaged in the development of sequence-based method for predicting PPIs, and the preliminary results have demonstrated their feasibility. Specifically, Bock and Gough [10] tried to solve this problem by using a support vector machine (SVM) with several structural and physiochemical descriptors. Martin et al. [11] used a descriptor called signature product, which is a product of subsequences and an expansion of the signature descriptor from chemical information to predict PPIs. Nanni and Lumini [14] proposed a method to predict PPIs based on an ensemble of K-local hyperplane distance nearest neighbor classifiers, where each classifier is trained using a different physicochemical property of the amino acids. Shen et al. [8] developed a SVM model by combining a conjoint triad feature with S-kernel function of protein pairs to predict PPI network and yielded a high prediction accuracy of 83.93%. Guo et al. [9] proposed a sequence-based method by combining auto covariance descriptor with SVM, and when applied to predicting yeast PPIs, it achieved very promising prediction accuracy. In our previous study, we also obtained good prediction results by using correlation coefficient [15] and autocorrelation descriptor [16], respectively.

In this study, we present a sequence-based approach for the prediction of interacting protein pairs using support vector machine (SVM) combined with local descriptors [17, 18]. The utilization of the local descriptors provides us with a chance to mine interaction information from the continuous and discontinuous amino acids segments at the same time [17]. The effectiveness of local descriptors depends largely on the correct selection of amino acid grouping [18]. By grouping amino aids into a reduced alphabet, we can create a more accurate protein sequence representation. Here, we adopted the amino acids grouping according to the successful use of classification in [8]. To evaluate the performance, the proposed method was applied to *Saccharomyces cerevisiae* and *Helicobacter pylori datasets.* Empirical results have shown that our SVM prediction model with local descriptors yields good performance. We also evaluated the performance of our method by preparing four cross-species data as the independent test set, which further demonstrates the effectiveness of our method.

2 Methods

2.1 Data Set

We evaluated our method on publicly available S.cerevisiae dataset, which were extracted from S.cerevisiae core subset of database of interacting proteins (DIP) [19] by Guo et al [9]. After the protein pairs which contain a protein with fewer than 50 residues or have ≥40% sequence identity were removed, the remaining 5594 protein pairs comprise the final positive data set. The non-interacting pairs were generated

from pairs of proteins whose subcellular localizations are different. The final data set consists of 11188 protein pairs, where half are from the positive data set and half from the negative data set. Three-fifths of the protein pairs which from the positive and negative data set were respectively randomly chosen as the training set, and the remaining two-fifths were used as the test set.

2.2 Local Protein Sequence Descriptors

To predict PPIs from sequences, one of the main computational challenges is to find the way to fully encode the important information content of proteins [11, 14]. In this study, each protein sequence is represented by local description of amino acid sequence [17, 18], and the PPI pair is characterized by concatenating the local descriptors of two proteins in this protein pair. To reduce the complexity inherent in the representation of the twenty standard amino acids and suit synonymous mutation, the amino acids were clustered into seven functional groups [8] based on the dipoles and volumes of the side chains (Table 1). Then the local protein descriptors abstract the features of protein pair based on the classification of amino acids. The process of generating local descriptors is described as follows.

Table 1. Division of amino acids based on the dipoles and volumes of the side chains

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group7
A, G, V	C	D, E	F, I, L, P	H, N, Q, W	K, R	M, S, T, Y

Firstly, for each protein sequence, every amino acid is replaced by the index depending on its grouping. For example, protein sequence AVDCNLSK is replaced by 11325476 based on this classification of amino acids. Secondly, we split the amino

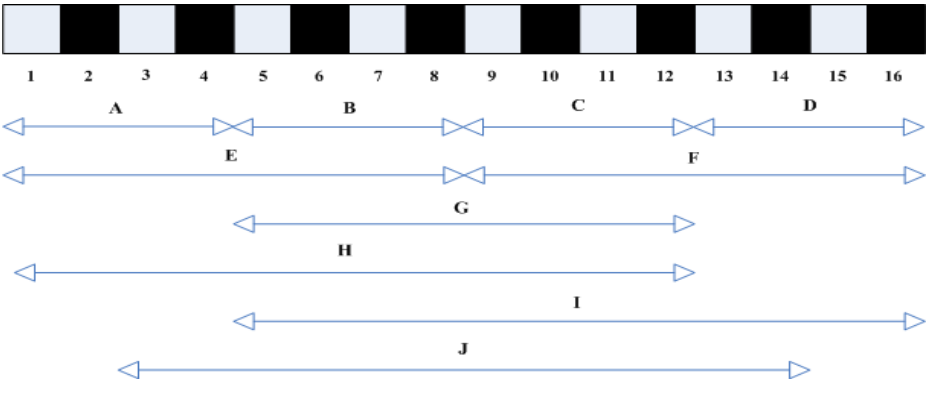


Fig. 1. Schematic diagram for constructing ten descriptor regions (A-J) for a hypothetical protein sequence. Adapted from Tong et al. [17] and Davies et al. [18]. The regions (A–D) and (E–F) are respectively generated by dividing the whole sequence into four equal regions and two equal regions. The region G, H, I and J stand for the central 50%, the first 75%, the final 75% and the central 75% of the entire sequence, respectively.

acid sequences into ten local regions of varying length and composition to describe multiple overlapping continuous and discontinuous interaction patterns within a protein sequence (see Figure 1). For each local region, three local descriptors, composition (C), transition (T) and distribution (D), are calculated. C stands for the composition of each amino acid group along a local region. T represents the percentage frequency with which amino acid in one group is followed by amino acid in another group. D characterizes the distribution pattern along the entire region by measuring the location of the first, 25, 50, 75 and 100% of residues of a given group.

For detailed descriptions of these descriptors, please refer to [17, 18]. Given that the amino acids are divided into seven groups in this instance, the calculation of the C, T and D descriptors generates 63 attributes in each local region (7 for C, 21 for T and 35 for D). The descriptors for all local regions were combined, resulting in 630 features representing the general characteristics of the protein sequence. Thus, a 1260-dimensional vector has been built to represent each protein pair and used as a feature vector for input into SVM.

2.3 SVM Optimization and Evaluation of Performance

The classification model for predicting PPIs was based on SVM. As a binary classification algorithm, SVM separates a given set of binary labelled training data (-1 and +1, in our case, non-binding and binding protein pairs) with a hyper-plane that is maximally distant from them (known as the maximal margin hyper-plane). The hyper-plane found by the SVM in feature space corresponds to a nonlinear decision plane in the input space. Each of the feature vector generated from the protein pair in the negative and positive dataset is assigned with a corresponding label of {-1} and {+1} respectively, indicating whether the pair is interacting with each other or not. The advantage of SVM is that there is no need to compute the coordinates of the data in the feature space, but instead simply computing the inner products between all pairs of data. This operation is often computationally cheaper than the explicit computation of the coordinates.

The LIBSVM package (http://www.csie.ntu.edu.te/~cjlin/libsvm) was employed in this work to do classification. A radial basis function (RBF) was selected as the kernel function. Two parameters, the regularization parameter C and the kernel parameters γ were optimized using a grid search approach. The prediction performance was evaluated by the overall prediction accuracy (ACC), sensitivity (SN), precision (PE) and Matthews correlation coefficient (MCC) [20]:

$$ACC=\frac{TP+TN}{TP+FP+TN+FN}, \tag{1}$$

$$SN=\frac{TP}{TP+FN}, \tag{2}$$

$$PE=\frac{TP}{TP+FP}, \tag{3}$$

$$MCC=\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FN)\times(TN+FP)\times(TP+FP)\times(TN+FN)}} \tag{4}$$

where TP, TN, FP and FN denote true positive, true negative, false positive and false negative, respectively. In addition, we also used the receiver operating characteristic (ROC) curve [21] to assess the prediction performance. An ROC curve is a graphical plot of the true positive rate (TPR) versus the false positive rate (FPR) for a binary classifier system as its discrimination threshold is varied. To summarize ROC curve in a single quantity, the area under an ROC curve (AUC) is used. The AUC score ranges from 0 to 1. When the AUC value of a predictor is larger than the area of other ROC curves, the predictor is regarded as a better one.

3 Results and Discussion

3.1 Assessment of Prediction Ability

In order to achieve good experimental results, the corresponding parameters for SVM were firstly optimized. Here, two parameters, C and γ were optimized using a grid search method within a limited range. Considering the numerous samples used in this work, 5-fold cross-validation was used to investigate the training set, which can minimize the overfitting of the prediction model. To test the robustness of the prediction model, five training sets and five test sets were prepared as described by the sampling method in Methods. Thus five models were generated for the five sets of data. The prediction results of SVM prediction models with local description of protein sequence are shown in Table 2. For all five models, the precisions are $\geq$88.66%, the sensitivities are $\geq$87.00%, and the prediction accuracies are $\geq$88.07%. On average, our method yields a PPI prediction model with an accuracy of 88.56±0.33%. To better investigate the practical prediction ability of our model, we also calculated the MCC and AUC values. From table 2, we can see that our method gives good prediction performance with an average AUC score of 95.07% and a MCC value of 77.15%. Further, it can also be seen in the experiments that the standard deviation of sensitivity, precision, accuracy, MCC and AUC are as low as 0.22, 0.60, 0.33, 0.68 and 0.39% respectively. The results illustrate that our model is an accurate and robust method for the prediction of PPIs.

Table 2. Prediction results of the test sets

Test set	SN (%)	PE (%)	ACC (%)	MCC (%)	AUC (%)
1	87.62	89.71	88.78	77.59	95.01
2	87.31	88.66	88.07	76.15	95.16
3	87.00	89.64	88.47	76.98	94.67
4	87.35	90.43	89.05	78.15	95.69
5	87.58	89.09	88.43	76.87	94.82
Average	87.37±0.22	89.50±0.60	88.56±0.33	77.15±0.68	95.07±0.39

There are two possible reasons that our SVM prediction model with local descriptors yields good performance. One is that the twenty standard amino acids have been divided into seven groups according to their related physicochemical properties of electrostatic and hydrophobic interactions[8]. Accordingly, the reduced dimension of vector space of protein sequence may partially overcome the overfitting problem [18]. In addition, the reduced but informative alphabet likely includes the information of synonymous mutations [8] for PPI because of similar characteristics within the same amino acid group. The other is that we used a series of local descriptors of varying length and composition to describe the physicochemical properties of proteins. Local descriptors account for the interactions between sequentially distant but spatially close amino acid residues [17]. As a result, such novel representation of local description of amino acid sequence enables our model to adequately capture multiple overlapping continuous and discontinuous binding patterns within a protein sequence.

3.2 Performance on Independent Dataset

As our method produced a good performance on the PPI data of S.cerevisia, we switched to evaluate the practical prediction ability of our final model against an independent dataset. Firstly, we constructed our final prediction model using the whole dataset (11188 protein pairs) with the optimal parameters (C = 32, γ = 0.03125). And then the prediction performance of the final predictor was evaluated using another dataset which is independent of the training dataset. Our model was trained on the S.cerevisia core subset in the DIP database; therefore we chose the other four species in this database as our independent test dataset. The performance of our method in predicting such samples is summarized in Table 3. The prediction performance in Caenorhabditis elegans, Escherichia coli, Homo sapiens, and Mus musculus achieved by our method is 75.73%, 71.24%, 76.27% and 76.68% respectively. It shows that the meta model can correctly predict the interacting pairs of three species with the accuracy of over 75% while the E. coli subset have a relatively lower accuracy which still >71%. It demonstrates that the SVM prediction model with local descriptors is able to achieve better performance towardscross-species dataset. We selected the PPIs data of S.cerevisiae to construct the final prediction model, so this model should represent the features of S.cerevisiae PPIs. At the same time, our model can also represent the features of C. elegans, E. coli, H. sapiens, and M. musculus, which is implied by the generalization ability of our model on these four species. Our findings indicate that our model may be applied to other organisms for which experimental data regarding PPIs may not be available.

Table 3. Prediction results on four species based on our model

Species	Test pairs	ACC (%)
C. elegans	4013	75.73
E. coli	6954	71.24
H. sapiens	1412	76.27
M. musculus	313	76.68

Interestingly, we found that there are some relationship between the prediction accuracies and the evolution of organisms. For example, S.cerevisiae and E. coli protein-protein interactions are not very closely related. That is to say, many proteins in the E. coli dataset were not presented in the S.cerevisiae dataset, and vice versa. As a result, when predicting E. coli from yeast, we had only limited success, as can be seen by the relative poor results reported in Table 3. On the other hand, if there is a close relation between species such as the yeast and M. musculus, our method can generate very promising results (see Table 3). These results agree well with the findings of Martin et al. [11]. At the same time, it should be pointed out that many PPIs in these four organisms dataset were not obtained by high-throughput proteome-wide methods and were small (for example, there were 1412 H. sapiens and 313 M. musculus interactions in the dataset). So it can be expected that with the increasing number of PPIs data, our model may not work as well.

3.3 Comparison with Other Methods

Many methods have been used in the prediction of PPIs. To compare prediction ability of the SVM prediction model using local descriptors with the existing methods, dataset H.pylori was constructed. The H.pylori dataset is comprised of 2916 protein pairs (1458 interacting pair and 1458 non-interacting pairs) as described by Maritin et al.[11]. Table 4 gives the average prediction results of 10-foldcross-validation over six different methods [10-14] on the H.pylori dataset. The methods of Bock and Gough ([10]), Martin et al. ([11]) and Nanni ([13]) are based on single classifer system to infer PPIs, while the methods of Nanni ([12]), Nanni and Lumini([14]) belong to ensemble classifier-based approach. From Table 4, we can see that the model based on SVM with local description of amino acid sequence gives good results with the average sensitivity, precision and accuracy of 0.851, 0.833 and 0.842, respectively. The results illustrate that our method outperforms other single classifier-based methods such as signature product method. It has pointed out that one single classification system cannot always provide high classification accuracy [22]. Instead, a multiple classifier system is proved to be more accurate and robust than an excellent single classifier [22]. However, it is remarkable that our prediction model obtain performance similar to those obtained by ensemble classifier-based methods. All these results demonstrate that the SVM classifier combined with local descriptors can improve the prediction accuracy compared with current state-of-the-art methods.

Table 4. Comparison of state-of-the-art methods on the *H.pylori* dataset

Mehods	SN	PE	ACC
Bock and Gough ([10])	0.698	0.802	0.758
Martin et al. ([11])	0.799	0.857	0.834
Nanni ([12])	0.806	0.851	0.83
Nanni ([13])	0.86	0.84	0.84
Nanni and Lumini([14])	0.867	0.85	0.866
Our method	0.851	0.833	0.842

4 Conclusions

In this paper, we present a simple and elegant sequenced-based approach to solve protein interaction problem. One particular feature of protein interaction is that the interactions usually occur in the discontinuous regions in the protein sequence, where distant residues are brought into spatial proximity by protein folding. However, litter literature tries to make use of such information. In the current study, a novel representation of local protein sequence descriptors was used to involve the information of interactions between distant amino acids in the sequence. A protein sequence was characterized by 10 local descriptors of varying length and composition. So this method is capable of capturing multiple overlapping continuous and discontinuous binding patterns within a protein sequence. As expected, experimental results show that our SVM-based predictive model with this encoding schemeis an important complementary method for PPI prediction.

We believe that the results can be further improved in the ways explained below. For example, the performance of local descriptors should be better enhanced if we optimize the amino acid grouping, i.e. we could extract more useful information by identifying the most efficient grouping. In addition, not all local descriptors are effective in the prediction. Some are less relevant to the prediction and some are redundant. We expect to improve the prediction accuracy by using some feature selection strategy, instead of using the whole local description features. Finally, it is evident that one single classification system cannot always provide high classification accuracy. Instead, a multiple classifier system is proved to be more accurate and robust than an excellent single classifier in many fields. Hence, using a multiple classifier learning approach could further improve the prediction accuracy.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. This research was supported by the grants of the National Science Foundation of China (Nos. 60472111 & 30570368).

References

1. Zhao, X., Wang, R., Chen, L.: Uncovering signal transduction networks from high-throughput data by integer linear programming. Nucleic Acids Research (36), 48 (2008)
2. Zhao, X., Wang, R., Chen, L., et al.: Gene function prediction using labelled and unlabeled data. BMC Bioinformatics 957 (2008)
3. Zhao, X., Wang, R., Chen, L., et al.: Protein function prediction with high-throughput data. Amino Acids 35, 517–530 (2008)
4. Ito, T., Chiba, T., Ozawa, R., et al.: A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. USA 98, 4569–4574 (2001)
5. Ho, Y., Gruhler, A., Heilbut, A., et al.: Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature 415, 180–183 (2002)
6. Zhu, H., Bilgin, M., Bangham, R., Hall, D., et al.: Global analysis of protein activities using proteome chips. Science 293, 2101–2105 (2001)
7. Skrabanek, L., Saini, H., Bader, G., Enright, A.: Computational prediction of protein–protein interactions. Molecular Biotechnology 38, 1–17 (2008)

8. Shen, J., Zhang, J., Luo, X., et al.: Predicting protein-protein interactions based only on sequences information. Proc. Natl. Acad. Sci. USA 104, 4337–4341 (2007)
9. Guo, Y., Yu, L., Wen, Z., Li, M.: Using support vector machine combined with auto covariance to predict protein-protein interactions from protein sequences. Nucleic Acids Res. 36, 3025–3030 (2008)
10. Bock, J.R., Gough, D.A.: Whole-proteome interaction mining. Bioinformatics 19, 125–134 (2003)
11. Martin, S., Roe, D., Faulon, J.L.: Predicting protein-protein interactions using signature products. Bioinformatics 21, 218–226 (2005)
12. Nanni, L.: Fusion of classifiers for predicting protein-protein interactions. Neurocomputing 68, 289–296 (2005)
13. Nanni, L.: Hyperplanes for predicting protein-protein interactions. Neurocomputing 69, 257–263 (2005)
14. Nanni, L., Lumini, A.: An ensemble of K-local hyperplanes for predicting protein-protein interactions. Bioinformatics 22, 1207–1210 (2006)
15. Shi, M., Xia, J., Li, X., Huang, D.: Predicting protein-protein interactions from sequence using correlation coefficient and high-quality interaction dataset. Amino Acids (2009), doi:10.1007/s00726-009-0295-y
16. Xia, J., Liu, K.H., Huang, D.: Sequence-based prediction of protein-protein interactions by means of rotation forest and autocorrelation descriptor. Protein and Peptide Letters (2009) (in press)
17. Tong, J., Tammi, M.: Prediction of protein allergenicity using local description of amino acid sequence. Frontiers in Bioscience: A Journal and Virtual Library 13, 6072 (2008)
18. Davies, M., Secker, A., Freitas, A., Clark, E.: Optimizing amino acid groupings for GPCR classification. Bioinformatics 24, 1980–1986 (2008)
19. Xenarios, I., Salwinski, L., Duan, X.J., Higney, P., Kim, S.M., Eisenberg, D.: The Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. Nucleic Acids Res. 30, 303–305 (2002)
20. Matthews, B.: Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim. Biophys. Acta 405, 442–451 (1975)
21. Zweig, M., Campbell, G.: Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin. Chem. 39, 561–577 (1993)
22. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms. Wiley, New York (2004)

Finite Element Analysis of Mould Temperature Change When Releasing Agent Sprayed

Xinping Hu[1,2,*] and Juan Zhang[1,2]

[1] School of Mechanical engineering, Shandong Polytechnic University, Jinan, Shandong, 250353, China
[2] Shandong Provincial Key Laboratory of Advanced Machine-making of Light Industry and Monitoring Technology, Jinan, Shandong, 250353, China
{huxxpp,juanzi0603}@163.com

Abstract. In order to introduce the analysis model basing on the actual situation of production to study the temperature change of mould, the mould and the spraying process were simplified. The distributions of temperature in the surface and different distances from the surface were studied in spraying conditions by the finite element analysis. The results showed that in the course of spraying, the surface temperature would decline sharply, but the inner temperature didn't change significantly. After three model comparisons, a 2D simple model was adopted to carry out the temperature change analysis, and the results were credible.

Keywords: Analysis models; mould; finite element analysis; temperature changes.

1 Introduction

With the method of the finite element analysis, the temperatures of e die casting moulds can be analyzed and the results can be used to the practical production. Under the die-casting process, thermal fatigue is the main factor that affects the life of mould. When the mould is sprayed releasing agents, its temperature changes seriously and the mould withstands instantaneous thermal stress [1]. So, understanding the mould temperature distribution is beneficial for prolonging the mould life [2].

According to the basic principles of finite element method in temperature analyzing, the First Law of Thermodynamics should be obeyed [3, 4]. According to the law of conservation of energy, the following equation could be obtained.

$$\rho c\left(\frac{\partial T}{\partial t}+V_X\frac{\partial T}{\partial X}+V_Y\frac{\partial T}{\partial Y}+V_Z\frac{\partial T}{\partial Z}\right)=q+\frac{\partial}{\partial X}\left(K_X\frac{\partial T}{\partial X}\right)+\frac{\partial}{\partial Y}\left(K_Y\frac{\partial T}{\partial Y}\right)+\frac{\partial}{\partial Z}\left(K_Z\frac{\partial T}{\partial Z}\right)$$

* **Foundation item:** Project(J09LD11) supported by Shandong Province Higher Educational Science and Technology Program; Project (BS2009ZZ010) supported by Shandong Province Outstanding Research Award Fund for Young Scientists.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 263–268, 2011.

where:
t stands for time;
T stands for temperature;
ρ stands for density;
c stands for specific heat;
q stands for heat flux density vector;
Kx, Ky, Kz stand for thermal conductivities in X, Y, Z directions respectively;
Vx, Vy, Vz stand for velocity in X, Y, Z directions respectively.

2 Analysis Models

The process of die casting and die spraying must be simplified and the analysis model should be created firstly according to actual production condition. In our study here, a 650-tons die casting machine was taken into consideration as an example to typical analysis. And the dimensions of the die were that fixed cavity depth was 50 mm, its insert thickness was 100mm, its rear section chamfered 25mm × 20mm and the four corners rounded R10mm. Therefore, the analysis model of stationary side mould was shown in Fig.1 and the model shown here was meshed.

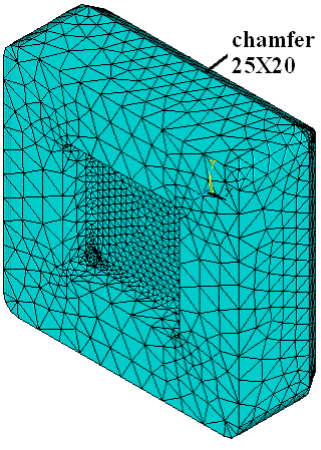
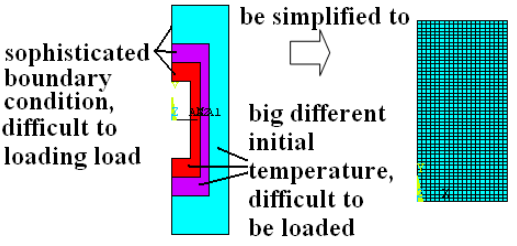


Fig. 1. Diagram of meshed model of stationary insert, it included 516121 nodes and 35789 units

The temperature of shape forming part of mould was 450 ℃, and the non-forming part was 250℃. Mould insert material was H13, its physical parameters were: specific heat was 489.9J / (kg^2·℃), thermal conductivity was 30 W / (m^2•℃), Young's Modulus was 210GPa, Poisson Ratio was 0.3, density was 7760kg / m^3.

In model shown in Fig.1, although the overall situation of the mould inserts could be reflected totally, the computer took too long time, it was difficult to load loading and the results were difficult to be treated afterwards. Therefore, the model must be further simplified.

A simplified thermal analysis model was shown in Fig. 2a. The external environment was sophisticated, and internal temperatures of the mould insert were difficult to be loaded. And non-cavity section of the mould should be considered in different way and its temperatures in different spots were different in die casting process, it was also difficult to be loaded. However, the initial and boundary conditions in this model were difficult to be defined. In our study, the surface temperature of cavity in the mould side and the gradient temperature change were the main and important information we want to get, so a much more simple model could be created shown in Fig. 2b. Considering the dimensions of the mould, its length was 200mm and width was 100mm.



a) difficult 2D model b) simplified 2D model

Fig. 2. Diagram of simplified model

Some comparisons of the above model in characteristics and applicability were shown in Table 1.

3 Analysis and Results

For the 650-tons die casting machine, the stationary mould spraying time is about one minute to three seconds. Here, the spraying time was set to two seconds. Adopting the model of Fig.2 b, analyzed according to following steps:

1. Entering the environment of thermal analysis.
2. Defining the thermal analysis unit type as PLANE55.
3. Determining the physical parameters of various materials.
4. Creating analysis models and dividing them into grids.
5. Setting the transient thermal analysis and selecting the method of "full".
6. Imposing a variety of loads.

On our experience of previous analysis, when the die releasing agent was sprayed on surface of the mould, its temperature was close to the boiling point of water. So, the external environment of the mould surface could be set 100℃. Based on the experience and relevant literatures, heat transfer intensity between the mould surface and environment could be set 10000W m^{-2} K^{-1}. When the casting was removed, the temperature of the mould surface was about 450℃ and the temperature of the peripheral block was about 250℃(in aluminum die casting, the temperature of the mould was about 200-240℃), and the initial temperatures of the mould in different spots were set according to the temperature-distance graph shown in Fig. 3.

Table 1. Comparisons of created models

model in Fig.1	model in Fig.2a	model in Fig.2b
1. High requirement of computer hardware; the result file could be several G	1. Low requirement of computer hardware; the result file could be dozens of M	1. Low requirement of computer hardware; the result file could be several M
2. Tens of thousands of equations to be solved, 20-30 hours to be calculated (CPU 2.8G 2.8G)	2. Hundreds of equations to be solved, a few minutes to be calculated (CPU 2.8G 2.8G)	2. Hundreds of equations to be solved, just a few minutes to be calculated (CPU 2.8G 2.8G)
3. Analyzing situations most totally in various parts of the mold and obtaining the most comprehensive information	3. Analyzing situations totally about most parts of the mold and obtaining the more comprehensive information	3. Analyzing situations about parts of the mold and obtaining the Specific information
4. Very difficult initial condi-tions, very complicate boun-dary conditions	4. Difficult initial conditions, complicated boundary con-ditions	4. Easy initial conditions, simple boundary conditions
5. Hard post-processing, difficult to observe the results	5. Simple post-processing, easy to observe the results	5. Simple post-processing, easy to observe the results
6.Unsuitable for cycle analysis	6. Difficult to cycle analysis	7. suitable for cycle analysis

Note: Solving equations, result file, analysis time are highly related to the meshing density of model.

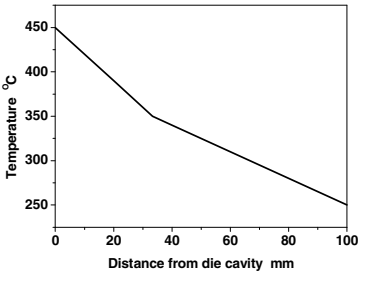


Fig. 3. Diagram of initial temperature of stationary insert

7. Setting the analysis result data output "each analysis sub-step". It produced a corresponding result. In order to analyze the temperature change with time, each result should be written to the result file.

8. Setting the analysis time and the time step size. The load step was set “Stepped” and the automatic time step was set “ON”. According to the preceding analysis, the analysis time was 2s and the time step was 0.01s.

9. Solving

10. Getting the solutions and analyzing them.

Some analysis results were showed in the Fig. 4 and Fig. 5.

From Fig. 4 we could see that, because of the releasing agent taking a lot of heat away, the surface temperature of the mould dropped sharply and with the spraying time became long, the surface and high temperatures declined. But, the temperature in the mould changed little because of the short diffusion time (0.5, 2s). The highest temperature was about 10mm away from the mould surface. So the forward of temperature gradient changed in this point: toward the inside cavity, there existed a positive temperature gradient and toward the outside a negative temperature gradient. This was affected by the transfer conditions. The bigger the temperature gradient was, the greater the mould thermal stress was. Greater thermal stress was bad for prolong mould life. So, reducing the mould temperature gradient would have a positive impact on the mould life.

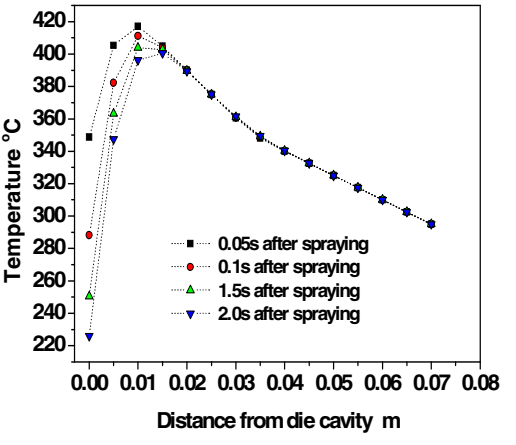


Fig. 4. Temperature distribution below die surface

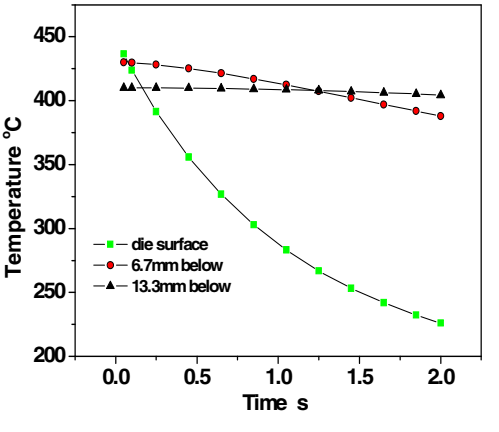


Fig. 5. Temperature history below die surface

Fig. 5 showed that the rapid temperature change of the mould only occurred in the surface within a certain range. When the distance from the mould cavity was about 13.3mm, the mould temperature changed very slowly.

4 Conclusion

This article described several models and compared their characteristics and applications. The distributions of temperature in the surface and different distances from the surface were studied in spraying conditions by the finite element analysis. The results showed that in the course of spraying, the surface temperature would decline sharply, but the inner temperature didn't change significantly. After three model comparisons, a 2D simple model was adopted to carry out the temperature change analysis, and the results were credible.

References

1. Arif, A.F.M., Sheikh, A.K., Qamar, S.Z.: A study of die failure mechanisms in aluminum extrusion. J. Mater. Process. Technol. 134, 318–328 (2003)
2. Hu, X., Zhao, G., Wu, B.: Temperature change and stress distribution analysis of die surface in Al-alloy die casting process and experimental study on die heat check. China Foundry 5, 47–53 (2008)
3. Prosenjit, S., Vijay, B., Tata, R.: Thermal structural analysis of SST-1 vacuum vessel and cryostat assembly using ANSYS. Fusion Engineering and Design 84, 7–11 (2009)
4. Hyder, M.J., Asif, M.: Optimization of location and size of opening in a pressure vessel cylinder using ANSYS. Engineering Failure Analysis 15, 1–19 (2008)

The Analysis on Applicability of Ride Comfort Standards to Vehicles

Aihua Tang[1], Chen Bao[2], Ou Jian[2], Guohong Deng[2], and Jianping Tian[1]

[1] No.180,Xueyuan Street, Huixing Road, Zigong, Sichuan.PRC
[2] Chongqing University of Technology.PRC
tahme@163.com, cbyztm@sina.com,
{oujian,dengguohong}@cqut.edu.cn, tjp893@126.com

Abstract. Aiming for the shortage of ride comfort standards appled to various vehicles and different terrains at present, evaluation systems applied to wheeled vehicles were introduced, which include of two categories. One system, which includes ISO 2631 and other similar standards, is generally used to evaluate ordinary vehicles. The other based on two limits, such as 6 W average absorbed power and vertical acceleration peak, is used by the United States of America and by North Atlantic Treaty Organization (NATO) to evaluate the NATO Reference Mobility Model (NRMM) .A passenger car test was carried out according to the national criteria: GB/T 4970 - 1996 and the objective ride comfort value was obtained. The applicability of evaluation systems based on limits was researched. The significance of this research is better to understand and apply the ride comfort standards to various vehicles and different terrains.

Keywords: vehicle, ride comfort, criteria, evaluation method.

1 Introduction

Four methods to evaluate ride comfort (human response to vibration) objectively are used throughout the world at present. The ISO 2631 standard is adopted mainly in Europe and the British Standard-BS 6841 is used in the United Kingdom. Germany and Austria use VDI 2057 while Average Absorbed Power (AAP) is used by the United States of America (USA)and by NATO in the NATO Reference Mobility Model (NRMM). There is a need to determine which of these four standards or methods are the best to evaluate ride comfort appropriately on various vehicles and different terrains, with the emphasis on ride comfort over predominantly rough, off-road terrains. The relationships between the measures of the standards can also be very useful when vehicles from manufacturers using different standards must be compared against each other, or to compare measurements with historical data[1].

2 Evaluation of Ride Comfort

The methods to evaluate ride comfort objectively of ordinary vehicles are different from military tactical vehicles. The following methods should be considered

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 269–275, 2011.

synthetically: ISO 2631 and similar standards applied to ordinary vehicles running on the road. The evaluation methods based on limits are for the off-road vehicles travelling on off-road terrains. A short description of the three objective methods used will now be given.

2.1 ISO 2631 (1997)

The revision of ISO 2631 released in 1997 [2] introduces new experience and research results. The standard rules that the Root Mean Square (RMS) value method can be applied to evaluate the human comfort and healthy effect of vibration, when the vibration peak factor is below 9 (peak factor is the ratio of root acceleration peak value to RMS value). The method applied to all kinds of ordinal vehicles, including all terrain vehicles travelling on the road. Vibration Dose Value (VDV) also can be used to evaluate as the vibration peak factor is higher than 9. Measurements which are for motion sickness range between 0.1 and 0.5 Hz in the ISO 2631 standard.

2.2 BS 6841

The BS 6841 standard considers a frequency range of 0.5–80 Hz. Instead of time dependency curves, a new method based on the concept of VDV was introduced. The frequency weighting for z-axis seat vibration is modified to be in closer agreement with the results of experimental research in this method. For each axis, a component ride value can be determined as well as an overall ride value. At first, the acceleration sampled is weighted. Different weight functions for three directions exist. As for the ride comfort, the RMS value of the weighted signal is determined as follows[3]:

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N} a^2{}_n} \tag{1}$$

The RMS value is compared with the subjective value to get an estimate of the most probable human reaction related to the vibration. The frequency range is extended to 0.5 Hz to compensate for measurements of motion sickness. The ride value obtained may be compared directly with that obtained at the same speed on the road in another vehicle. A doubling in the ride value corresponds to a doubling of the vibration discomfort.

In principle, the methodology and calculation of the standard are the same as the ISO 2631-1:1997 standard. The main difference between the BS 6841 standard and the ISO 2631 (1997) is that weighting replaces Wk where Wb is the result of intensive laboratory studies. Fig. 1 gives a comparison of acceleration weighting curves used in ISO 2631 (1997) and BS 6841 (1987) respectively.

2.3 VDI 2057

In 1963, the Society of German Engineers published VDI 2057 standard making them the first to produce a standard to quantify ride comfort. In principle, the VDI standard defines a calculated ride comfort index (K-factor) that is compared with a subjective table to determine the ride as subjectively experienced by humans. In 1979, the VDI

standard adopted the tolerance curves of the ISO 2631 (1978) but kept the K-factor for the subjective comparison of the sensation perceived by humans. The acceleration data is converted into the frequency domain using a Fast Fourier transform (FFT). From the FFT, the RMS value at the third octave centre frequencies is determined. These result in single value at the centre frequencies. The RMS acceleration data is weighted and K-values for the z-direction are determined as follows.

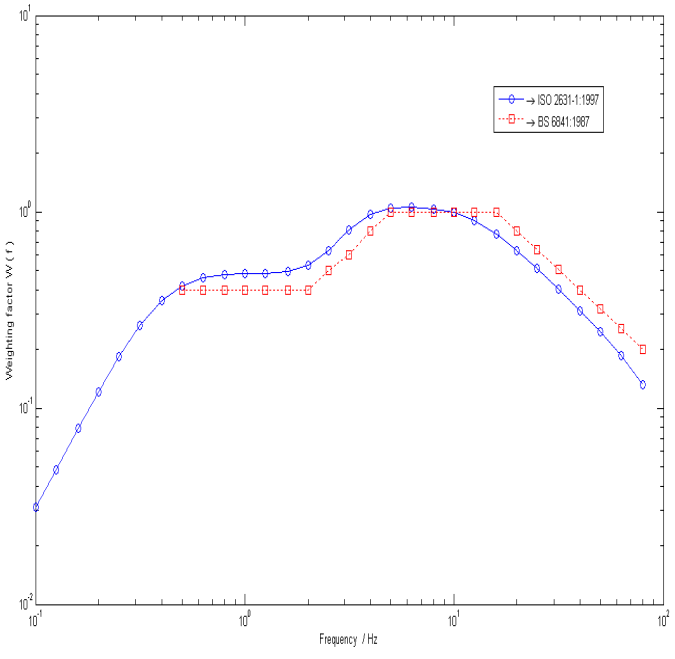


Fig. 1. Comparison between acceleration weighting curves for BS 6841:1987 and ISO 2631-1:1997

$$\begin{aligned} 1 \le f \le 4Hz \quad & K_z = 10 \cdot a_z \cdot \sqrt{f} \\ 4 \le f \le 8Hz \quad & K_z = 20 \cdot a_z \\ 8 \le f \le 80Hz \quad & K_z = 160 \cdot a_z \cdot f \end{aligned} \tag{2}$$

The weighted signal is then plotted against limit curves. In principle, these are the same as the limit curves of ISO 2631-1:1997. The frequency bandwidth ranges between 1 and 80 Hz[4], [5]:

3 The Evaluation Methods Based on Limits

Average absorbed power was developed by the US Army Tank-Automotive Command in 1966. Studies have shown that the human body behaves in an elastic fashion. Under vibration, the body's elasticity produces restoring forces that are

related to displacement. This process continues until the energy imparted is dissipated or removed. The time rate of energy absorption is referred to as the power. The power can be computed in the frequency domain as well as in the time domain[6], [7].

$$AAP_t = \sum_{i=0}^{N} K_{f_i} a_{f_i}{}^2 \tag{3}$$

Frequency weighting ranges between 1 and 80 Hz. Below 1 Hz the method is not successful. The AAP weighting curve strongly emphasis the visceral resonance around 4–5 Hz presumable because most energy is absorbed in these softer tissues.

4 Road Test and Evaluation of Ride Comfort

4.1 Road Test

According to the national standard—GB/T 4970-1996[8], a road test of a passenger car was carried out. The car was driven in the asphalt road surface which equivalent to B level. The test speeds were 40, 50, 60, 70, 80km/h. The speed was maintained constant by driving the vehicle in a selected gear during the test through the steady speed section. Two tri-axial seatpad accelerometers (six channels) were used to measure lateral, longitudinal and vertical acceleration on different positions. These measurements were used to determine objective ride comfort value on seats in the vehicle. The data was acquired later. The value was determined finally by measuring acceleration on the passenger seat and left rear seat in the vehicle.

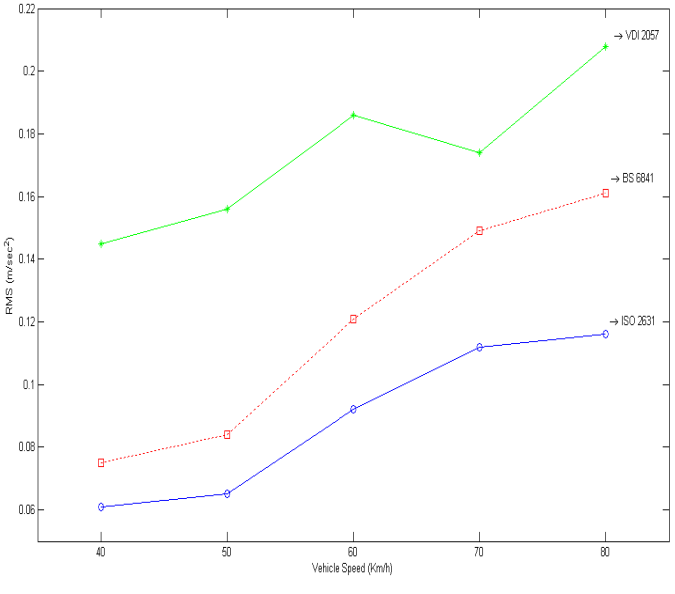


Fig. 2. Total weighted root mean square speed characters for the driver seat

4.2 Determination of Objective Evaluation Value

The evaluation value can be determined by measuring the acceleration of four points in five different vehicle speeds. Initially, five tests were performed where acceleration was measured on the vehicle seats. Ride comfort value was determined according to ISO 2631-1:1997 and BS 6841:1987.

Fig. 2 shows there is a visible difference existing in the curve between VDI2057 and others. The curves of ISO 2631 and BS 6841 are closer. Besides, the trends for ISO 2631 and BS 6841 are in very good agreement. Furthermore, the evaluation value is consistent with the actual situation of passenger cars.

The weighted RMS value for 2 points in typical conditions is summarized in table 1. It can be found from table1 that the RMS value in the x and y directions is smaller than the z-direction. The z value at the rear is significantly higher than that at the left rear, while the driver and rear x and y value were smaller. It is clear that the evaluation value of ride comfort is dominated by the vertical acceleration value.

Table 1. ISO 2631-1:1997 weighted RMS value for 3rd gear 70km/h

Position	RMS	Total weighted RMS
X Seat front	0.086	—
Y Seat front	0.045	—
Z Seat front	0.188	0.231
X Seat rear	0.094	—
Y Seat rear	0.048	—
Z Seat rear	0.146	0.208

4.3 Limiting Value for Ride Comfort

(1) Limits for ISO 2631 and similar standards

In this research, a comparison of ISO 2631 and BS 6841 guidelines compared to the subjective comments was summarized in Table.2.

(2) 6W-based Average Absorbed Power Limits

In a tactical test, the tri-axial acceleration of seat bearing surface was measured to calculate AAP. An algebraic sum of tri-axial AAP was the total average absorbed power. The total AAP is under 6W, when the test vehicle passed the stipulative test road according to 6W-based AAP Limit value. For the experiment indicated, as the total average absorbed power surpasses 6W, the passenger will focus on holding the handrail tightly. Assuming experience provided by some researches exploring the relaship between average absorbed power and RMS, the 6W of AAP is equivalent to the RMS value of 2.07 $m \cdot s^{-2}$.

(3) 2.5g-based limiting value of vertical acceleration peak value

A group with a different height bump obstacles surfaced road was used to evaluate shock resistance. The section of a bump obstacles surfaced road is semicircular. The

height of bump ranges from 4 to 12 inches. In actual test, the peak value of vertical acceleration for the driver seat bearing surface were acquired. The limiting value of peak value requires the peak value of vertical acceleration less than 2.5g while the test vehicle passing through the stimulative bump obstacles. In the military tactical vehicle's test, the special absorption dynamometer was applied to measure AAP and peak value of vertical acceleration by the U.S arm.

Table 2. ISO 2631 and BS 6841 guidelines compared to the subjective comments

W-RMS $/m.s^{-2}$	Weighted Vibration level $/dB$	Subjective Comment
<0.315	110	Not uncomfortable
0.315~0.63	110~116	A little uncomfortable
0.5~1.0	114~120	Fairly uncomfortable
0.8~1.6	118~124	Uncomfortable
1.25~2.5	112~128	Very uncomfortable
>2.0	126	Extremely uncomfortable

5 Conclusions

1）For the vehicles of different types and purposes (Such as ordinary vehicles and military tactical vehicles, etc.), the vehicle speed and roads commonly used are different, which should be considered while carrying out the road test. Therefore, the selection of evaluation standard should keep consistent with the vehicle speed and road. 2）6W-based Average Absorbed Power Limit value and 2. 5g-based limiting value of vertical acceleration peak value evaluation methods are the results of U.S troops implemented a large number of experimental verification of military tactical vehicles. 3）ISO 2631 and similar standards can be utilized in the ordinary vehicles and military tactical vehicles travelling on the road. Military tactical vehicles can use evaluation methods based on Average Absorbed Power and limiting value of vertical acceleration peak value to evaluate the ride comfort objectively.

Acknowledgements

This work was financially supported by the Sichuang University of Science & Engineering Foundation (2010XJKRL013),Talent Introduction Program of Sichuang University of Science & Engineering. Educational Reform Program of Sichuang University of Science & Engineering (JG-1010) and Natural Science Foundation (CSTC,2008BB6337).

References

1. Els, P.S.: Journal of Terramechanics. J. Mol. Biol. 42, 47–64 (2005)
2. Mechanical vibration and shock – Evaluation of human exposure to whole-body vibration, Part 1:General requirements, ISO 2631-1, 2nd edn. The International Organisation for Standardisation (July 15, 1997); P.S. Els / Journal of Terramechanics 42, 47–64 (2005)
3. British Standard Guide to measurement and evaluation of human exposure to whole body mechanical vibration and repeated shock, BS 6841. British Standards Institution (1987)
4. Human exposure to mechanical vibrations whole-body vibration. Verein Deutcher Ingenieure, VDI 2057 (September 2002)
5. Hohl, G.H.: Ride comfort of off-road vehicles. In: Proceedings of the 8th International Conference of the ISTVS, Cambridge, England, August 5–11, vol. I(III) (1984)
6. Scarlett, A.J., Price, J.S., Stayner, R.M.: Whole-body vibration: Evaluation of emission and exposure levels arising from agricultural tractors. In: Proceedings of the 9th European Conference of the ISTVS, Harper Adams, UK, September 8–11 (2003)
7. Pradko, F., Lee, R.A.: Vibration comfort criteria. SAE Technical Paper 660139. Society of Automotive Engineers, Warrendale (1996)
8. GB/T 4970-1996. Method of random input running test-Automotive ride comfort

The Research and Realization of Power Supply in High-Speed DSP System

Zhang jun-xing, Bo chun-juan, and Shi li-xin

Dalian Nationalities University
116600 Dalian, China
zhangjunxing@dlnu.eud.cn

Abstract. Power supply is a difficulty in high speed electronic system design. In the process of data acquisition and processing with high-speed DSP, according to the diversity of the power supply and the complexity of the system signals, the high-speed circuit board's power supply design rules and power supply filtering methods are studied and the corresponding design rule are put forward. The Monitoring Circuit of Power Supply and Power Supply Decoupling methods are discussed. The main rules should be attention in PCB layout and design are presented. System debugging results show that the power system is stable, reliable and can meet the electromagnetic compatibility requirements.

Keywords: DSP; Power Supply; PCB design.

1 Introduction

TMS320DM642 is a high-speed data acquisition and processing chip which developed by Texas Instruments (TI), with performance of up to 5760 million instructions per second (MIPS) at a clock rate of 720 MHz and it's external memory interface can be clocked up to 133M/S. The chip can execute up to 5.7 billion times Multiplication and Accumulation (MAC) operation per second, so the data processing platform which builds with TMS320DM642 have powerful data processing capability. Fig.1 shows the data processing system which design based on the chip.

The system is mainly consisted of TMS320DM642, A/D, D/A, DRAM, FLASH, power, clock, reset and the JTAG debug interface and other components. Where: two 4 Meg×32bit chips consist SDRAM, so the total capacity is 4 Meg×64bits and the maximum bus clock is 166MHz. FLASH total capacity: 4Mx 8-bits. The conversion rate of A/D and D/A up to a maximum of 27MHz, in the aim to meet the need of video data capture and output. During the system design process, the external data bus and address bus clock is 133 MHz [1].[1]

2 System Power Requirements and Design Plan

The core voltage of TMS320DM642is 1.4V, I/O voltage is 3.3V. The A/D converter is TVP5150, a chip especially suitable for video signal acquisition for it can do both

[1] Supported by Fundamental Research Funds for Central Universities(Grant No.DC10020112).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 276–281, 2011.

A/D conversion and video decoding. The core and the analog input voltage is 1.8V, I/O voltage is 3.3V. Other peripheral chips' voltage is 3.3V, and the power supply voltage of the system is 5V. The system power demand is shown in Table 1. In order to provide ideal power supply to the system, reasonable voltage chip is necessary. At present the most commonly used voltage regulator are LDO voltage regulator and switching regulator. Their performance comparison is shown in Table 2.

There are 1.4V, 1.8V, 3.3V, 5V four voltage levels in this system. Since TMS320DM642 dissipate more power, the rated current of core and I/O is 890mA and 210mA.Due to the chip's high frequency and low sensitivity to power supply noise, so we choose switching regulator. When design the system, we Select two DC-DC power supply chips TPS54310 to provide 1.4V (core voltage) and 3.3V (I/O voltage) for TMS320DM642. While the 3.3V I/O voltage also doubles as the system power supply, load external components. The input voltage of TPS54310 is 3-6V, output current up to 3A, can satisfy system needs completely.

For high-speed A/D,D /A and analog signals in this system, because of theirs very sensitive to power supply noise[2], we choose low-noise LDO power supply chip TPS76733 and TPS76718 as the power conversion chip. Their rated current is 1A, which can satisfy the need of A/ D and D/A. The power supply of I/O of the A/D, SDRAM, FLASH and Ethernet equipment is 3.3V, which are supplied directly by the system power. Fig.2 shows the specific power management.

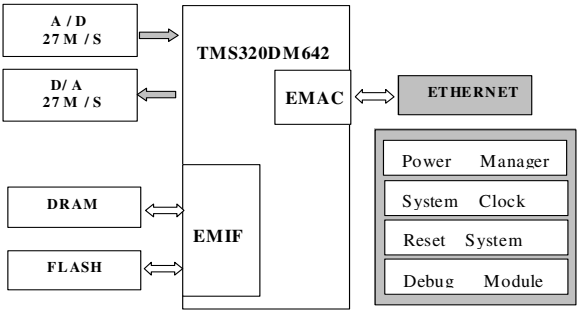


Fig. 1. System Structure Schematic

Table 1. System Power Requirement

Required Supply Voltage	Voltage (V)
System Power	5
DSP core voltage	1.4
DSP I /O voltage	3.3
A/D, I/O voltage	3.3
A/D digital, analog voltage	1.8
D /A digital, analog voltage	3.3
Ethernet digital, analog voltage	3.3
Other devices	3.3

Table 2. LDO and Switching Power

LDO Regulator	Switching Regulator
conversion efficiency low	Typical conversion efficiency of 92%
Fast response	Slow response
Low Noise	High Switching Noise
Decoupling Capacitors, May Cause System Instability	Decoupling capacitor less impact on the system
Low-cost	Relatively high cost

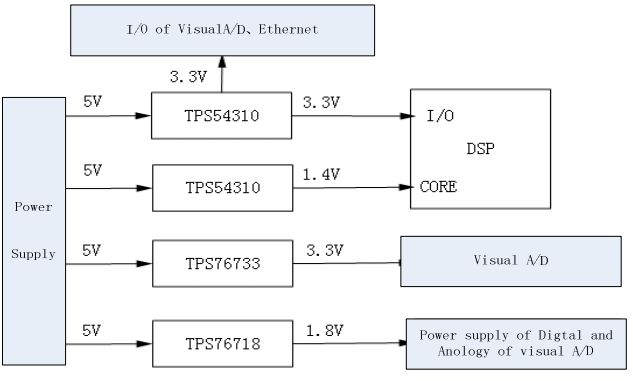


Fig. 2. System Power Management Solution

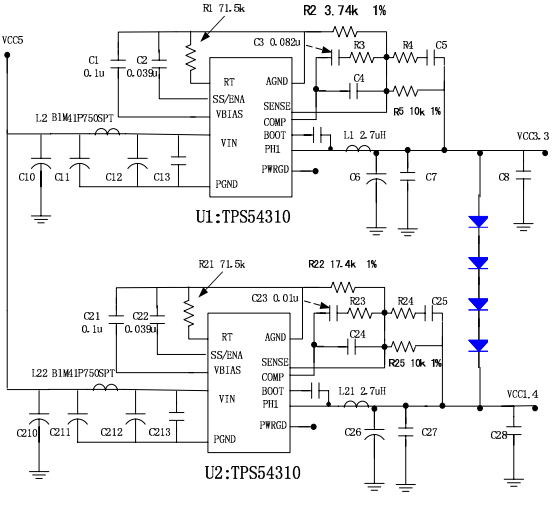


Fig. 3. CPU core and I / O power supply

We adopt power chip TPS54310, which produced by Texas Instruments, in this system. Its input voltage range between 3V and 6V, can provide current up to 3A, output voltage can adjust between 0.9V and 3.3V. Because TPS54310 is only a single power adjustable voltage output, two TPS54310 are needed to achieve 3.3V and 1.4V, the voltage requirements of TMS320DM642. Fig. 3 shows the circuit connection.

U_1 provide the required system voltage of 3.3V, U_2 generate 1.4V voltage. U_1 in the R_2 and R_5 constitute a feedback loop to determine the system's output voltage.

The formula is: $Vout= V_{Ref} \times R_5/ R_2$.

Vref=0.891V, the output voltage:Vout=0.891x10/3.74=3.3V

Similarly, the output voltage of U_2 is 1.4V. Where, L_1, C_6, C_7 form the output filter circuit. We take magnetic bead as L_2 to make power can filter high frequency, which meet the needs of the system to suppress EMI.

3 Monitoring Circuit of Power Supply

TMS320DM642 consume more power. In order to ensure the stable and reliable work of system, it is necessary to monitor the working conditions of power supply. We use TPS3307-25D in the system. It can provide voltage monitoring and reset functions for the circuit, and can simultaneously monitor the working conditions of the three voltages. Table 3 shows the monitoring voltage level and threshold voltage [3].

Table 3. Monitoring and Threshold Voltage

TPS330	Monitoring voltage (V)			Threshold voltage(V)		
SENSEn	SENSE 1	SENSE2	SENSE3	SENS E1	SENSE2	SENSE3
Voltage(V)	3.3	2.5	User	2.93	2.28	1.25

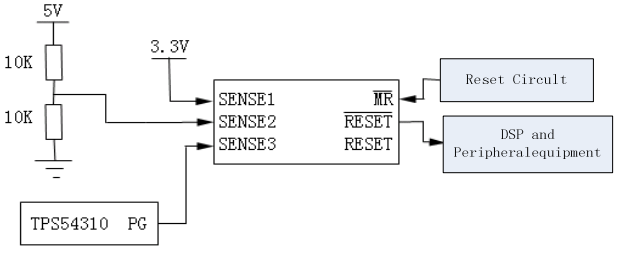


Fig. 4. Power Monitoring and Reset

When the power is higher than 1.1V, the voltage monitoring circuit start monitoring the input of SENSEn. As long as the monitoring voltage SENSEn less than the reset threshold voltage VIT+, RESET# will always reset effective. An internal time delay control the time that output delay active (low) change into the inactive state (high) to ensure the reliability reset of the system. Delay time is fixed at

200ms, and begin at the time that SENSEn higher than its threshold voltage (VIT +). When we monitor SENSEn land below its threshold voltage (VIT-), RESET# output active (low).

For this system, when monitoring 3.3V, 5V and the TPS54310 power status, we should make 3.3,5 V partial pressure of 2.5V, connect TPS54310 chip pin PG (power conversion complete signal) to SENSE1, SENSE2, SENSE3 to monitoring the state of 3.3V, 5V, and power chip TPS54310. Reset circuit, connect to pin MR#, and provide power on reset and manual reset. Schematic diagram shown in Fig. 4.

4 Circuit Board Design

Printed circuit board (PCB) is support pieces of electronic components and devices in electronic products. It provides electrical connections between components and devices. The quality of PCB has a great impact on product performance [4].

4.1 PCB Stack Strategy

Power line and ground is the most important part on printed circuit board. In order to reduce interference among power, ground and signal, and meet the needs of EMC, power supply layer, ground layer and signal layer should be separate on multilayer alignment board [5]. After analysis and simulation, we use 6-layer, layer distribution and defined as follows. The minimum alignment width and line spacing is 5mi1, size of through hole is 10/18mi1, material is FR4, PCB board thickness is 1.6mm. The structure shown in Fig.5.

Layer 1 Top - the main component side
Layer 2 Ground Plane - Formation
Layer 3 Inner - Signal Layer
Layer 4 Inner - Signal Layer
Layer 5 Power Plane - power layer
Layer6 Bottom - second component side

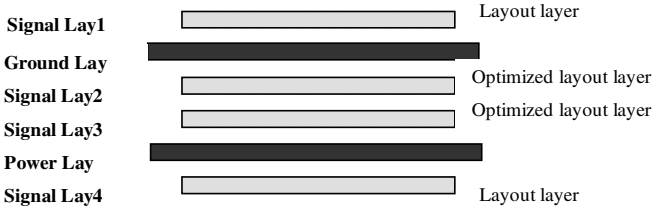


Fig. 5. PCB Layer Structure Diagram

4.2 Power Supply Decoupling

In the system, TMS320DM642 core and I/O switching enjoy a high speed, we should ensure that the surge of its chip core and I/O voltage were less than: 50mV and 20mV. Thus, a good chip decoupling design is indispensable. During the design process, we take several measures in the following [6]:

- To increase the number of decoupling capacitors: The best design is add decoupling capacitors between each power and ground. However, the number of power pin and ground pin are more than 200 (548 pins totally), so we can only appropriately increase the number of decoupling capacitors on the condition that the wiring allow.
- Add 8 10UF Tantalum capacitor as low-frequency filter, 4 of them work as core power of the chip and the other 4 work as I/O power supply.
- In order to reduce the equivalent inductance and equivalent resistance of capacitor, we take small surface-mount chip package and try to shorten the distance between chips.

5 Conclusions

In this paper, we analyze the requirements of High-speed DSP System; the system is represented by TMS320DM642. What's more, high-speed circuit board's power supply design rules and power supply filtering methods are studied and the corresponding design rules are put forward. System debugging results show that the power system is stable, reliable and can meet the electromagnetic compatibility requirements.

References

1. Schell, Tsividis, B., Yannis: A continuous-time ADC/DSP/DAC system with no clock and with activity dependent power dissipation. IEEE Journal of Solid-State Circuits 43(11), 2472–2481 (2008)
2. Byun, Kwon, K., Young-Su: Digital audio effect system-on-a-chip based on embedded DSP core Source. ETRI Journal 31(6), 732–740 (2009)
3. Sharma, R., Chakravarty, T.: Analytical model for optimum signal integrity in PCB. Interconnects using ground tracks. IEEE Transactions on Electromagnetic Compatibility 51(1), 67–77 (2009)
4. Ricchiuti, V.: Power bus signal integrity improvement and EMI mitigation on multilayer high-speed digital PCBs with embedded capacitance. IEEE Transactions on Mobile Computing 2(4), 314–321 (2003)
5. Derryberry, G.: Realistic implementation of guidelines signals integrity screening for PCB designers. Electronics World 116(1886), 22–25 (2010)
6. Chen, W., Yao, T.: Research on PCB simulating design and analysis based on signal integrity (SI) theory. Journal of Wuhan University of Technology (Transportation Science and Engineering) 29(2), 273–276 (2005)

Study on the Competence Evaluation Index System of College Counselor

Yong Zhang[1,2] and Guohong Jiang[2]

[1] Control Engineering Department, Chengdu University of Information Technology, No.24, Block 1, Xuefu Road, Chengdu 610225 P.R. China
scuzhangyong@126.com
[2] School of Public Administration, Sichuan University, Chengdu, 610064, P.R. China

Abstract. College counselors are a crucial part of college faculty, and also the key members of ideological work. But there are many limitations in the standards of traditional counselor's selection and evaluation, which have influenced counselor's selection and appointment. Based on competency theories and models, this paper combines with the working characteristic of college counselors and sets up a competence evaluation system of College counselors. In it, we are determined four first level indicators, includes knowledge, skill, quality and individual motivation. We also determined the other 18 secondary indicators such as general knowledge and knowledge of education and psychology. Then the weights of evaluation index are determined by virtue of the Analytic Hierarchy Process (AHP).Thus construct the college counselors of competence evaluation model and provide the reference for establishing a scientific and reasonable evaluation system of counselors.

Keywords: college counselor; competence; Analytic hierarchy process; applications.

1 Introduction

College counselors are a crucial part of college faculty, who shoulders the responsibilities of conducting moral education and providing guidance for undergraduates' healthy growth. Due to the complex and complicated tasks, heavy workload without any spatial or temporal boundary and the posteriority and longevity in the effect upon every undergraduate, it is hard to evaluate college counselor's work by rational and conventional means. Therefore, an appropriate way to evaluate the job performance counts for much in both motivation and construction of college counselors. So far, many colleges have already established evaluation systems for college counselors, in which the deficiencies upon rationality and scientificity lead to the failure of accurate evaluation. Moreover, as researches go deeper into the theory of competence, some models, which mainly consist of questionnaire and interview surveys, have been already applied to the systems. However, these models are seldom put into practice and have no weight determined in every aspect that matters to the job performance differently. Compared with the traditional methods to examine and

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 282–289, 2011.

evaluate a college counselor's job performance, the theory-based one agrees with the variety of counseling work, the flexibility of managing methods and the acuity of political thoughts. Besides, it dynamically reflects the core of a college counselor's responsibility. In this paper, the theory of competence is combined with analytic hierarchy process to establish a new evaluation system that contains weights in all aspects regarding competence for college counselors.

2 Determination of Competence Evaluation Index of College Counselor

Since McClelland published the article "Testing for Competency rather Than Intelligence" which proposed the concept of competence for the first time in 1973, different scholars put forward different definitions. At present, the definition of competence proposed by McClelland and Spenser in 1994 was generally accepted. They believe: "competence can be motive, traits, skill, self cognition, attitude or values, specific knowledge, cognitive or behavioral skills – which can be certain individual characteristics measured and calculated accurately." These features can clearly differentiate excellent performance from low performance.

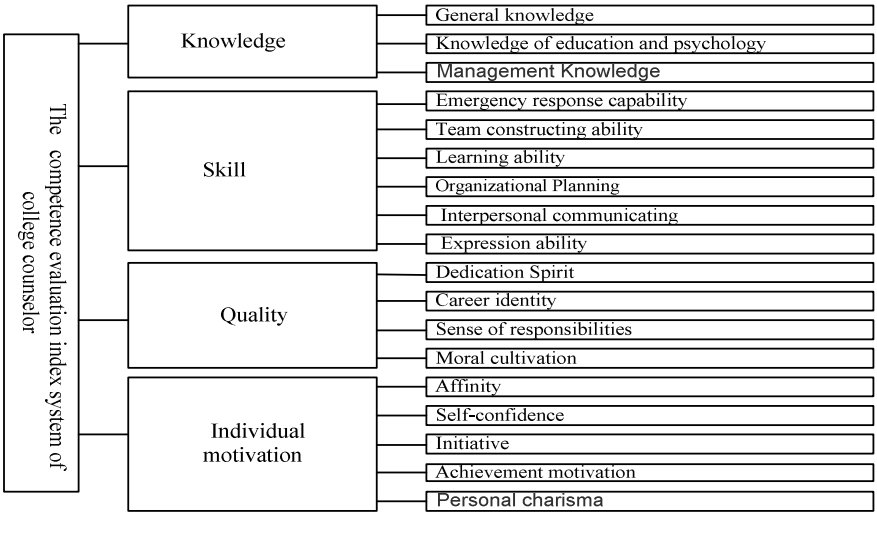
The researchers employ questionnaire and interview methods to study the competence of counselors and reach the different research results. Please see the table 1 below.

Table 1. Research results

Researchers	Research Results
Guqian, Yang jiping[1]	16 competence characters include verbal expression skills, communication skills, strain capacity, organization skills, caring for students, innovation ability, observing ability, professional loyalty, personal charm, promoting the students development ability, ideological and moral cultivation, psychological counseling ability, introspection cognitive ability, understanding and respecting the students, principles, participation ability, in which college students believe that the most significant five competence characteristics in order are: understanding and respecting the students, verbal expression skills, principles, caring for students, professional loyalty.
Zhang qin Su wenming[2]	Competence character model contain ten basic competence characteristics in three categories. Three categories are service and management skills, service and management quality, and ego growth. Service and management skills includes the innovation ability, communication and coordination, problem solving, planning organization; Service and management includes service consciousness, professional loyalty, initiative and the professional quality; The ego growth includes continuous learning and psychological adjustment
Chen yansong[3]	College counselor competency model includes three dimensions, namely the professional attitudes dimensions: eager for success, initiative, service consciousness; Knowledge-skill dimensions: professional knowledge, learning ability, analysis ability, interpersonal understanding, cultivating others, team leadership, communication ability; personality charm dimensions: self-confidence, introspection ability, affinity
Hao yingjie[4]	College counselors' competence CCR model was constructed from the three aspects of capacity, quality, and personality charm. Capabilities include learning ability, working ability, leadership ability and innovation ability; Quality includes the ideological and moral qualities, scientific and cultural quality, and physical health quality; Personality charm includes enjoyable work, loving and caring, affinity, evaluating students correctly, handling principle of fair, just and open, learning to listen, strong sensitivity, overall situation consciousness and noble moral sentiment

College counselor is the person who engaged in student's ideological and political education in universities of China. Their basic task is to endeavor every efforts to educate students in both politics and ideology according to the training goal and the law of development under the leadership of the party organization in universities. Depend on the results of the studies above and with the concern of the college counselor's capability, this essay has selected 30 indicators, which as follows: general knowledge, knowledge of education and psychology, expression ability, interpersonal communicating, organizing ability, team constructing ability, learning ability, observing ability, innovation ability, psychological adjusting ability, introspection cognitive ability, executive ability, emergency response capability, moral conduct and personal integrity, responsibility, career identity, spirit of devotion, initiative, achievement motivation, personal charisma, self-confidence and affinity. Combining the interview results between good and general counselors, eventually we sorted a competence evaluation index system of college counselor. There are 4 indicators in the first-level: knowledge, skill, quality and individual motivation. From within to outward, the external surface is knowledge and ability, quality is the middle-tier and the individual motivation as the core layer. Knowledge indicators include: management knowledge, knowledge of education and psychology, general knowledge; skill index includes: expression ability, interpersonal communicating, organizational planning, team constructing ability, learning ability, emergency response capability; quality indicators include: moral cultivation, sense of responsibility, career identity, dedication spirit; individual motivation indicators include: initiative, achievement motivation, personal charisma, self-confidence and affinity. Thus, we can get the competence evaluation index system of college counselor.

Table 2. The competence evaluation index system of college counselor



3 Analytic Hierarchy Process Methodology

There are many ways to calculate the weight of indicators. This paper establishes an index system which is a multi-level framework; So AHP will be useful in calculating the weights. The Analytic Hierarchy Process (AHP) was developed by T L Saaty. It is a new system analysis method as well as a combination of quantitative and qualitative method. The core of this theory is that the formalization of a complex problem using a hierarchical structure and it is a multi-criteria decision-making approach that employs pairwise comparisons (Saaty. 1994). [5]

There are three steps in AHP:

Step1: In order to achieve the goal, design of the decision is hierarchy. This is because problems can be divided it into different factors, According to the relationship between the factors, thus can form a multi-layer analysis structure model.
Step2: Make pairwise comparison. Each element in an upper level is used to compare the elements in the level immediately below with respect to it. (Quoted from Saaty, 2008a) [6]
Step3: Calculate the priorities and check the consistency of the judgments.

3.1 Pairwise Comparison

When making comparisons, we need a scale of numbers that indicates how many times more important or dominant one element is over another element with respect to the criterion or property with respect to which they are compared. Table 3 exhibits the scale. By making comparisons and scoring, we can get a comparison matrix which is based on the Saaty's scale as given in Table 3. When making pairwise comparison, we must repeat the answers——the ratio of the importance of two factors.

Table 3. The scale of AHP

Intensity of Importance	Description
1	Two activities contribute equally to the objective.
3	Experience and judgment slightly favor one activity over another.
5	Experience and judgment strongly favor one over another.
7	An activity is strongly favored and its dominance demonstrated in practice.
9	The evidence favoring one activity over another is of the highest possible order of affirmation.
2, 4, 6, 8	Intermediate values between the two adjacent judgments
X(i,j)=1/(xj,i)	If C is 8 times as good as D, then D is 1/8 times as good as C.

3.2 Calculate the Priorities

After construct the pairwise comparison matrix M=[mij]n×n, we first calculate the product of elements of each row in the matrix, and then calculate the arithmetic square root of the product. At last, the weights are computed. The procedures are as follows.

the product of elements of each row:

$$Mi = \prod_{j=1}^{n} a_{ij} \quad (i=1,2,\ldots,n);$$

arithmetic square root of the product：

Pi= $\sqrt[n]{Mi}$

Weight calculation:： Wi= $Pi \Big/ \sum_{i=1}^{n} Pi$ （ i=1,2,…,n ）；

3.3 Check the Consistency of the Judgments

When establishing the matrix, we use the pairwise comparison, which is a subjective assignment method. Objectivity is not strong enough. We have to judge whether there is a certain degree of consistency.

Thorough calculating the consistency index and ratio, we can evaluate the reliability of judgment matrix and decide whether to accept the judgment matrix. The procedures of consistency check are as follows:

1.Calculate the consistency index CI：CI= $(\lambda_{\max} - n)/(n-1)$

λmax is the maximum eigenvalue of judgment atrix.λmax= $(1/n)\left[\sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij} W_j \Big/ W_i)\right]$

2. Calculating the consistency ratio CR:

CR= CI/RI . RI is the average random consistency index obtained from a randomly generated pairwise comparison matrix. Table 4 shows the RI. If CR<0.1, the consistency of the judgment matrix is acceptable, or to adjust judgment matrix again.

Table 4. RI

N	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

4 The Weights of Evaluation Index are Determined by AHP

As the qualitative data obtained from academicians and college counselors, the Pairwise Comparison Matrix is framed using Saaty's scale. The Pairwise Comparison Matrix for the college counselors' competence evaluation considered in this study is given in Table 5.

Table 5. The pairwise comparison matrix of the competence evaluation index system

The pairwise comparison matrix of the competence evaluation index system of college counselor

	Knowledge	skill	quality	Individual motivation	Priorities
knowledge	1	1/2	1/3	1/4	0.11
skill	2	1	1/2	1/3	0.16
quality	3	2	1	1/2	0.27
Individual motivation	4	3	2	1	0.46

After calculating the priorities, we must check the consistency of the judgments. Take the table5 as an example.

$$\lambda\max=(1/n)\left[\sum_{i=1}^{n}(\sum_{j=1}^{n}a_{ij}W_j\Big/W_i)\right]=(1/n)\sum_{i=1}^{n}[(a_{i1}W_1+a_{i2}W_2+a_{i3}W_3+a_{i4}W_4)/W_i]$$

$=1/4\times(\ 3.5909+4.1771+4.2593+4.1739\ )=4.0503$

$CI=(\lambda_{max}-n)/(n-1)=(4.0503-1)/(4-1)\ =0.01677$.Finding the RI in the table, we can get RI equals 0.9, so $CR=CI/RI=0.01677/0.9=0.019<0.1$ the comparison is consistent.

Table 6. The Pairwise Comparison Matrix for knowledge

The Pairwise Comparison Matrix for knowledge				
	general knowledge	Knowledge of ducation and psychology	Management Knowledge	Priorities
general knowledge	1	1/2	1/3	0.16
knowledge of education and psychology	2	1	1/2	0.3
Management Knowledge	3	2	1	0.54

CR(knowledge)=0.008<0.1

Table 7. The Pairwise Comparison Matrix for skill

The Pairwise Comparison Matrix for skill							
	Emergency response capability	Team constructing ability	Learning ability	Organizational Planning	Interpersonal communicating	Expression ability	Priorities
Emergency response capability	1	1/2	1/3	1/5	1/6	1/7	0.04
Team constructing ability	2	1	1/2	1/3	1/4	1/5	0.06
learning ability	3	2	1	1/2	1/3	1/4	0.11
Organizational Planning	5	3	2	1	1/2	1/3	0.16
interpersonal communicating	6	4	3	2	1	1/2	0.25
expression ability	7	5	4	3	2	1	0.38

CR(skill)=0.019<0.1

Table 8. The Pairwise Comparison Matrix for quality

The Pairwise Comparison Matrix for quality					
	Dedication Spirit	career identity	Sense of responsibilities	moral cultivation	Priorities
Dedication Spirit	1	1/3	1/5	1/7	0.06
career identity	3	1	1/3	1/5	0.12
Sense of responsibilities	5	3	1	1/3	0.26
moral cultivation	7	5	3	1	0.56

CR (quality) =0.043<0.1

Each individual Pairwise Comparison Matrixis checked for consistency using the above procedure. The Pairwise Comparison Matrix based on the knowledge criterion is given in Table 6. The priority based on skill criterion is given in Table 7. The Pairwise

Comparison Matrix based on the quality criterion is given in Table 8. The Pairwise Comparison Matrix based on the individual motivationcriterion is given in Table 9.

All the CR of the Pairwise Comparison Matrices are less than 0.1, the Pairwise Comparison Matrices are acceptable. So we can construct the competence evaluation index system of college counselor. Table 10 shows the competence evaluation index system of college counselor.

Table 9. The Pairwise Comparison Matrix for Individual motivatio

The Pairwise Comparison Matrix for Individual motivation						
	affinity	personal charisma	achievement motivation	self-confidence confidence	initiative	Priorities
affinity	1	1/2	1/3	1/4	1/5	0.06
personal charisma	2	1	1/2	1/3	1/4	0.10
achievement motivation	3	2	1	1/2	1/3	0.16
self-confidence confidence	4	3	2	1	1/2	0.26
initiative	5	4	3	2	1	0.42

CR(Individual motivation)=0.015<0.1

Table 10. The competence evaluation index system of college counselor

	First level indicators	secondary indicators
the competence evaluation index system of college counselor	knowledge（0.11）	general knowledge 0.16 knowledge of education and psychology 0.3 Management Knowledge 0.54
	skill（ 0.16 ）	emergency response capability 0.04 team constructing ability0.06 learning ability 0.11 Organizational Planning 0.16 interpersonal communicating 0.25 expression ability 0.38
	quality（ 0.27 ）	Dedication Spirit 0.06 career identity 0.12 sense of responsibilities 0.26 moral cultivation 0.56 Affinity 0.06
	individual motivation (0.46)	personal charisma0.10 achievement motivation0.16 self-confidence confidence 0.26 initiative0.42

5 The Weights of Different Evaluators

The evaluation of competence college counselors comes from their own, superior, functional departments and their students. Their evaluation of their competence is different. So we can construct a pairwise comparison matrix. The pairwise comparison matrix is followed as table 11.

Table 11. The pairwise comparison matrix of different Evaluators

The pairwise comparison matrix of different Evaluators					
	Self	Departments	Superior	Students	Priorities
Self	1	1/3	1/4	1/5	0.07
Departments	3	1	1/2	1/3	0.17
Superior	4	2	1	1/2	0.29
Students	5	3	2	1	0.47

After calculating the priorities, we must check the consistency of the judgments. $\lambda max=4.0525$, $CR=0.019<0.1$. The consistency of the judgment matrix is acceptable.

To sum up, We can construct the competence evaluation index system of college counselor.

$$U=\sum_{j=1}^{4}(\sum_{i=1}^{4}Z_i \bullet W_i)\bullet C_j \quad Z_i=\sum_{t=1}^{n}Y_t \bullet W_t \quad i,j=1,2,3,4 \quad t=1,2,\cdots n$$

U represents the final score of Competency assessment. Zi represents the scores of each level indicators,Wi represents the weight of each level evaluation indicators, Cj represents the weight of the evaluators, Yt represents scores of secondary indicators, Wt represents the weight of secondary indicators in the level of evaluation indicators.

6 Conclusion

In this work, competence theory is introduced in the college counselors evaluation syst em, which lead to the new perspective for the evaluation and has been more theoretcal basis. Particularly, proposed competency model can improve and optimize the index s ystem of college counselors. However, competence theory is only one of many referen ces approach, so the design college counselors evaluation should be based on actual an d reasonable assessment model to adjust and make it more practical. Constructing scie ntific and reasonable college counselors competence evaluation system is a complex work. It need to understand the counselors job responsibilities, nature of work and clar ify the relationship between them to evaluate the counselors work fairly. AHP is used to determine the weight indicator system at all levels, by which the cousel-or compete ncy model is constructed to avoid the weight being determined at will in traditional me thods, so the credibility of the evaluation is improved.

References

1. Yang, J.-P., Gu, Q.: The Study of University Counselor Preliminary Competence. Journal of Shanxi University (Philosophy and Social Sciences) (6), 56–58 (2004)
2. Zhang, Q., Su, W.: College instructors in education management for model and its application. Education Exploration (3), 90–91 (2010)
3. Chen, Y.: Competency modeling of political instructorsin the university: An empirical study. School of Economics and Management (4), 23–26 (2010)
4. Hao, Y.: College instructors competence research. Journal of National Academy of Education Administration 49(6), 22–25 (2007)
5. Saaty, T.L.: How to Make a Decision. The Analytical Hierarchy Process, Interfaces 24(6), 19–43 (1994)
6. Saaty, T.L.: Decision making with the analytic hierarchy process. Int. J. Services Sciences 1(1), 83–98 (2008)

Computer Compose Principle Teaching and Developing Students' Research and Innovation Competence

Shangfu Hao[1], Wenliang Wang[2], Lili Huang[3], and Zhenghong Guo[1]

[1] School of Information Science and Engineering
Hebei North University,
Hebei, Zhangjiakou 075000, P.R. China
[2] Hebei North University Library,
Hebei, Zhangjiakou 075000, P.R. China
[3] Wuyi University,
Guangdong, Jiangmen 529000, P.R. China
zjkcxh5656@sina.com, zjkcxh5656@163.com
{yellowtiger999,hbnuwyq}@126.com

Abstract. The unique characters of computer compose principles teaching makes it a successful course in cultivating students' competence in doing research and creative work. This paper analyzes the advantages of this course in practical teaching and some basic principles for developing students' practical skills and promoting new ideas. It also explores some teaching strategies so as to better develop students' innovation spirits and to inspire original ideas.

Keywords: computer compose principles; research competence; innovation competence; teaching reform.

1 Introduction

Higher education has multi-requirements in developing students' capability in various forms. The development of students' research competence and new ideas is one of the essential parts in teaching, which is also an important action widely carried out in colleges and universities. Research competence refers to forming suspicious opinions, criticizing, and carrying out investigations. The learners should have the ability to apply what they have learnt, book knowledge and theories, to various kinds of practice, such as science, art, and technology, and hence come up with original ideas, unique techniques and new theories of economic value, social value and ecological value. Innovation competence consists of five parts: a sense of innovation, the foundation of innovation, the methods of innovation, the environment of innovation and the intelligence of innovation which including observation ability, thinking ability, imagination ability, and operation ability.

Developing students' innovation competence has long been promoted by education ministry in China, yet still a new field in teaching for both teachers and students. Problems and difficulties are sure to follow with the trail in teaching. Therefore, identifying the problems in the teaching process and finding ways to deal with the

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 290–296, 2011.

problems has profound meaning in theory and practice. It is also an important course to upgrade the integrated quality of teachers and students.

2 An Analysis of the Character of Computer Compose Principle Course and the Advantages of Developing Students Innovation Competence

Computer compose principle is a basic specialized course related to computer science and relevant specialties .The course tells about computer hardware and its work principles. By offering class on Von neumann machine's five parts---arithmetic unit, controller, storage unit, input device and output device, students will be familiar with structures of the five parts and their functions. They will also master the ways the computer works and the process and the control of computer internal data. The most typical character of the course is that it is theory-oriented, overloaded with up-to-date knowledge and covering a wide range of different fields.In addition, most students are overanxious for quick results while some focus on computer software, neglecting hardware. Therefore, the outcome of the course is that teachers find it hard for teaching while students complain about its difficulties in learning. For example, one of the learning targets of the course is master control's work principle. Without actual seeing or touching, students' understanding have to be based on their imagination. That is why most students will report that it is very hard for them to learn the course well. Yet, this part plays a key role in later learning of computer specialized courses. Therefore, exploring a better way of teaching, and an easy learning for students in order that students will be motivated in computer study and using new ideas has become our research project.

Computer compose principle course is the hardcore in computer related specialized courses in colleges and universities. With the help of multi-media assisted teaching, building network teaching platforms and virtual environment, the theory teaching and laboratory teaching will be very effective. In order to promote students' integrated application skills and innovation competence, the computer compose principle course provides classes on testing, designing and researching. It has its great advantage in theory teaching, laboratory teaching, engineering design teaching and promoting students innovation competence.

2.1 The Advantage of Computer Compose Principle Course in Developing Students' Innovation Competence

Computer compose principle course includes eight basic parts---computer system theory, arithmetic unit, operational method, storage unit, instruction system, central processing unit, bus system, peripheral apparatus, input and output devices. These parts will help students to understand the basic structure, basic principle and the basic design of a computer. The large quality of theory knowledge and systematic teaching and learning make it the advantageous course in developing students' creativity competence.

For instance, when learning micro-program control unit, students find it hard to understand how instruction works in CPU. In order to solve the problem, an new

teaching method is used : explaining step by step and forming three-dimensional concepts. In the form of time, fetch instruction event happens during the fetch time. In the form of space, the internal storage instructions flow to instruction register; the internal data flow to general purpose register. The second step is to make full use of software devices and designed models.Students are encourages to use FLASH and other devices to design the work flow of the instruction in CPU. Over ten years of trail teaching and learning has proved it an effective method. Students not only master the instruction work process but also increase their sense of creativity.

In the teaching, teachers find out that most students favor software more than hardware. However, as a computer specialized course, there should not be boundary between software and hardware. Hardware without software development will become an empty body without thoughts. Software without the support of hardware will become a deserted land without woods and water. The assignments on designing by using software replacing hardware will help students develop their creative thinking, innovative skill, and the sense of promoting new ideas.

2.2 The Advantage of Computer Compose Principle Laboratory Teaching in Developing Students' Innovation Competence

Computer compose principle laboratory teaching has a strong part in the overall computer teaching. It requires high quality and suitable experiment devices and advanced experiment techniques. In the actual practice, students are required to complete not only the assignments on the textbooks, but also the new testing designed by themselves. In so doing, students will find much pleasure in bringing on new ideas.

Over the years, our college has been using the experiment device produced by Qing Hua University, from tec-2 to tec-xp. The reason we use their device is that the device has the whole set of functions, because it is 8-bit machine is 16-bit machine as well. It has combined logic control and Mir program control. It also has its own monitor process and instruction system. All in all, it is a genuine top quality computer system. The weaknesses of tec-xp is its complicated techniques with more than 100 light-emitting diodes and its more than 100 jumper crowbar switches, which leaves less room for hardware expansion . That is one of the reasons that students find it difficult to learn the course well.Considering this, we incorporate multi-media teaching into laboratory teaching. We have also developed, by using flash and Dreamweaver, a set of laboratory teaching courseware to go with computer compose principle course. This courseware is easy to understand with its straightforward displays and strong dynamic characters.One the one hand, It helps to complete dynamic display and analog simulation, which are the basic laboratory requirements. On the other hand, the courseware emphasizes the students' actual skills in operation and application of theory to practice. What's more, students' self-study ability and design ability have been improved.

2.3 The Advantage of Engineering Design Teaching in Computer Compose Principle Course Displayed in Fostering Students' Innovation Competence

In order to make up the weaknesses in traditional computer compose principle laboratory teaching, to produce a large number of intellectuals of overall-skills in

system design, social application and innovation, we have carried out some new ways of teaching in laboratory teaching in computer compose principle course by adding creating new design experiment classes. The added experiment classes only set up the requirements for the experiment project. First, students have to complete a large portion of verification basic experiments. Second, they are supposed to have a full understanding of computer compose principle stated in the textbooks. Then, students should give the assignment a serious thinking using more than one kind of knowledge learnt. Finally, they will be able to produce their self-designed hardware circuit and software program. In the process of debugging hardware circuit, software program and correcting mistakes in design, students obtained dramatic progress in related knowledge, such as, the use of microprocessor internal source, the analysis of control timer, the connection of system bus and the application of peripheral interface chip.

We have also gradually built up open experimental environment (including open teaching hours, open teaching contents, open laboratory devices). Students therefore, are able to make their own arrangements according to their time, their interests, and their majors. In such a relaxing experiment environment, students will come up with their own experiment topic, make their own experiment plan and choose an appropriate flexible time to complete the experiments assignment.Teachers enthusiastically encourage students to speak up their own experiment ideas and framework, and try their best to provide guidance and support for students to carry out meaningful experimental study.The engineering design course offers students a platform to do wide imagination and bold creative work. Students are eager to produce new ideas and designs, which in turn motivate them to make new things and to use new techniques. This is the result the engineering design course in computer compose principle desires to achieve.

3 The Strategies in Computer Compose Principle Course to Develop Students' Innovation Competence

3.1 Careful Choice on Textbook and Better Organization of Teaching Contents

A textbook is a vehicle for teaching contents, teaching approach and knowledge. It is also an essential tool to fulfill teaching task. A good textbook ensures successful teaching and learning.Given students' characteristics and special features of this course, over the years, we have been using the textbook Computer Compose and Computer Mechanism, edited by Wang Aiying, and published by Qinghua University. So far, we have been using the textbook from the 2nd version to the latest 4th version. Its reasonable organization, stand-out key points, explanation the profound in simple terms and close connection between all knowledge points enable teachers to properly organize teaching contents.Following the principle of from the top to the bottom, from the external to the internal, the textbook displays to the students the mechanism and the principles of arithmetic unit, storage unit, control unit, input and output device.Meanwhile, the textbook reinforces the students' understanding of the whole set by going from the whole to the part and from the part to the whole.Besides, the textbook also incorporates teaching reform, quality education and innovation development education into course teaching. Therefore, a virtuous circle of course

construction has been built with the characters of course contents and course reform as its core part, quality teaching material as the lead, the teaching material building combined with teaching reform as its co-force to form a spiral development.

3.2 Imagery Teaching to Inspire Students

In the computer compose principle course, there exist many difficult concepts, technical terms and work principles of some important parts. We adopted real examples in life to make comparisons. Apart from that, we use multi-mediate teaching facilities to mock learning. We try hard to develop students' imagination so that they can convert abstract models into concrete real models.Take, for example, the execution of instruction in CPU. In terms of time, sequential control is used; while in space, parts are used to label the differences. Since students lack imagery comprehensive ability, they usually fail to form concrete models in their minds. In our teaching, we teachers ask each student to work out a set of motional pictures of execution of instruction in CPU by using flash or Dreamweaver. Via self-mock experiments, students are able to form clear pictures of what CPU models are, how the instruction execution is done and how micro-program control is carried out. A survey with many graduates over the years, this method has proved very effective in understanding work principles of the parts, inspiring students to active study and interests in learning computer principles.

In the teaching of computer compose principle, we apply elicitation method to inspire students' enthusiasm in active learning and active acquisition of knowledge.Computer compose principle course has the character of theory-orientation with abstract concepts. Pre-reading will help students to get easy access to the concepts and principles. Before attending the class, students should preview the textbook with questions and problems in mind. On finding the answers, more questions are to be expected. Then, when the students are having the actual class, the abstract concepts will not be very challenging.For example, when having a class about hard disc interface, questions are designed, such as what are the popular interfaces nowadays, why the transmission speed of SATA interface is faster than that of PATA, and, why PATA interface is no longer in use. These questions are for students to think about while doing group discussion or self study before class time. In class, students are encouraged to do presentations explaining the answers. Finally, teacher-based summary will follow.We all know well, a class is a place where interactions are carried out between teachers and students, between students and students, and between students and teachers. It is also a place questions and problems are answered or solved. What's more, it is a place where students acquire knowledge and build up knowledge base.

3.3 Strengthen Experiments to Develop Students' Practical Skills

The purpose of doing experiments is not only those students perform what is asked in the textbook, but also that students should be able to bring forth their original ideas and plans. Such action is designed to develop students' innovation and exploration spirit and to encourage them to probe into new world.

We have also gradually built up open experimental environment including open teaching hours, open teaching contents, open laboratory devices. Students therefore, are able to make their own arrangements according to their time, their interests, and their majors. In such a relaxing experiment environment, students will come up with their own ideas and plans for the experiments. On the same time, teachers enthusiastically encourage students to speak up and provide guidance and support for students to carry out meaningful experimental study. Such flexible yet well- guided experiment environment will surely lead to the increase of students' enthusiasm in active self-study and strengthen their sense of innovation.

3.4 Strengthen Research and Train Students' Development Skills

Supporting students in starting their own careers and inspiring their enthusiasm in developing new things has always presented an important section in promoting students' innovation spirit and social practice skills. Statistics show that 20% of our students have the desire to do research work and to develop new things. As teachers, we should cherish their enthusiasm with on-time guidance and encourage them to build up a sense of development. If possible, teachers can get students involved in teacher-oriented researches.A number of students in our college, under the guidance of teachers, completed the research on "Virtual Laboratory System Based on VC++ Computer Principles". The research project, using higher language, meets the requirements in laboratory teaching which is an indispensable part in computer compose principle course.During the research, students are required to not only intensively understand the operation principles of every part of a computer, but also apply theory to the function of experiment box and the design of laboratory teaching.It has turned out to be an effective way to motivate students' enthusiasm in scientific innovation, to promote students' development spirit and hand-work competence, and to strengthen students' interests in designing computer compose principle teaching. The new method has greatly inspired students to make their decisions on development topics in accordance with their own interests. Students' potential has been brought into full play, and their sense of innovation has also been developed.

3.5 Developing Network and Performing Shift-Mode Teaching

In the teaching of computer compose principle, most parts of basic components of computer, its work principle, the producing of sequential logic control signal, the time and space relation between instruction and data are explained in terms of logic diagrams. The large number of complicated diagrams has caused difficulties in class teaching. What's worse, some of the teaching contents cannot be put across to students by words and gestures. On the other hand, a fixed courseware can not meet the requirements of flexible teaching. It is also time-and-energy-wasting that every teacher work on their own courseware.In order to enable students to get easy access to the course and to meet the needs of teachers doing easy class preparing, making full use of our school's website system, we developed Computer Compose Principle based on the Web. It is a system emerging research with study.The system is the result of combined efforts of teachers and students. The software is designed and

produced according to software engineering principle and its method. During the research, students' active involvement is played to the most. What's more, what once seemed difficult to understand and too abstract to interpret now has become easy to grasp and clear in this courseware.

Making full use of network is another important means of improving traditional teaching. It is mutually complementary. In net-teaching, students learn to unload, on the teaching platform, the courseware, to submit assignments, do perform self-study and practice, to raise questions, and to provide suggestions, while teachers can upload courseware, assign after-class work, send answers, and provide solutions. The net-teaching and net-learning brings about changes in traditional teaching mode, improves net-interactions, and increases teaching and learning efficiency. It has gained better feedbacks from teachers and students as well.

4 Conclusion

During recent years, in accordance with the character of computer compose principle course, we have done a large amount of work on effective teaching of key points and difficult points, on developing students' practical skills, and on improving students' innovation competence. The reform in teaching has proved a success and the effect is obvious. Not only the teaching quality has steadily improved, with students excellent feedbacks, but also course construction has fulfilled. However, there also exist some problems, such as, the difficulty in grading students' self-study and experiment scores as more and more students are taking the course, the ever-increase of work hours for teachers, and the copying and cheating among some students in homework and experiments. The problems need further study and the situations require further improvement.

References

1. Wang, A.: Computer compose principle and structure, 4th edn. Qinghau University Press, Beijing
2. Ma, L., Zhang, Y., Song: Study and practice on developing computer majors' hardware-oriented innovation skills. J. Computer Education 2, 1–4 (2010)
3. Hao, S., Sun, B., Wang, Y.: Teaching Study on the Course Computer Organization Principle. J. Teacher Education Research 1, 81–82 (2009)

Mobile Phone Enabled Barcode Recognition for Preferences Monitoring

Eman Yasser Daraghmi, Chia-Feng Lin, and Shyan Ming Yuan

Distributed Computing System, Institute of Computer Science and Engineering,
National Chiao Tung University Hsinchu, Taiwan
{Eman.yasser85,teralin,smyuan}@gmail.com

Abstract. In this paper, we present an innovative software system called Preferences Monitoring that makes use of modern middleware technologies and available hardware and communication systems. People can use the system to easily buy products that fit with their preferences. The system consists of mobile-phone side to scan product barcodes, submit user inquiries to a server, and display results. Results are related to whether users can eat or drink the product based on their preferences such as religion (no meat, beef, pork, or wine), or based on health settings such as allergies, chronic diseases or calories. Also, the system includes web services to facilitate user access to online profiles for editing accounts, viewing history-shopping and subscribing to other services. The preliminary results of testing our system show that the system can provide useful services for consumers. We also highlight some technical issues that are related to reading barcode using mobile phones.

Keywords: Barcode recognition, consumer preferences, J2ME middleware.

1 Introduction

Consumer products packaging has introduced a noticeable amount of product-related information. This includes nutritional information, ingredients, country of origin, and contents which may cause allergy. There is also a plenty of additional product-related information that is not directly printed on the product packaging due to size constraints and possibly commercial considerations, e.g. reviews by consumer watch groups or price comparisons. For these reasons, the information cannot be customized for consumers who find it difficult sometimes to choose products that fit with their preferences.

In this paper, we address the problem of people shopping according to their preferences such as culture and health. People who travel to new countries usually face difficulties to adapt themselves to the new life. One of the most important difficulties is food/drink shopping especially when the language written on products is not the same as their native language. People will find it difficult to find the desired product, to know the ingredients of a product, and read warnings and other important information. Furthermore, people may come from different cultures with religious restrictions on food and drinks as people might not be allowed to eat meat, beef, or pork or drink wine. In addition to shopping in a foreign country, people in their native

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 297–302, 2011.

countries may need to monitor what they buy. People with chronic diseases or who have allergy to a specific food or drink need to watch carefully what they eat and drink. Also, people who are on diets need to calculate the number of calories in their food.

Nowadays, manufacturers and markets use barcode to track a product so that storing and selling become easier. A barcode is a small image of lines and spaces that represent numbers and other symbols. A barcode symbol typically consists of five parts: a quiet zone, a start character, data characters (including an optional check character), a stop character, and another quiet zone [1]. It represents data by combining these parts that containing products information. This information can be extracted using an optical scanner barcode reader.

Different standards have been used in the barcode systems. One of these standards is the EAN-13 barcode which is a 13 digit (12 + check digit) ant it is a superset of the original 12-digit Universal Product Code (UPC) system developed in the United States. The EAN-13 barcodes are used worldwide for marking products often sold at retail point of sale. The numbers encoded in EAN-13 bar codes are product identification numbers. All the numbers encoded in UPC and EAN barcodes are known as Global (GTIN).

The barcode system can be integrated with new technologies such as mobile phones enabling more useful services. Mobile phones comprise display, long-range communication capabilities, processing, and user profile storage capabilities [1]. Employing mobile phones and image recognition technologies to identify a product barcode in shopping offers simple and faster interaction. In this paper we present a system that enables consumers to know with a single click more information about products and find an answer for their question “Is this product fine for me?”

2 Related Work

In this section we review several systems developed based on barcode, mobile phones and image recognition technologies. Junaini and Abduallah developed a system to automatically scan barcode to recognize Halal food (food can be eaten by Muslims) using mobile phones [2]. Another system was proposed in [3] for visual impaired people. The system provides barcode scanning and text to speech conversion. A management system of grocery production based on barcode identification technology was proposed in [4]. The authors of [5] introduced an algorithm for 1D and 2D barcode using embedded digital signal processor (DSP).

The above studies stated some related problems to mobile-phone enabled barcode recognition. Consumers sometimes cannot capture a barcode precisely due to limited code image size, blur code image due to reflections or poor lighting conditions, or out-of focus mobile phone distance [6],[7]. Solutions to this problem vary between titling, skewing or rotating a barcode image [8]. Another solution is an algorithm proposed to rectify the distorted code image [9].

Our system is different from the previous systems in three aspects. The first aspect is using efficient barcode recognition that supports camera zoom-in and zoom-out functionality. This reduces the effect of the size and distance constraints. The second aspect is using middleware technologies that support web services to allow users to add any service they want and enable it on their mobile phones. Finally, we include

different types of communication technologies such as 3G and SMS. The 3G communication offers high data transfer speed and capacity. The SMS can be used if a mobile phone doesn't have 3G communication since SMS can provide useful information at real time [10].

3 System Overview

The system is called "Preferences Monitoring". It is based on EAN-13 barcode recognition. It allows users to predefine their profile from their account that contain user food culture, e.g. Christianity, Judaism, Islam, etc, the substances they are allergic to, calories based diet, user likes and dislikes. If a user is then holding a mobile phone in front of a certain product, e.g., in the supermarket, the phone displays a message like "This product is fine for you", "Warning, this is not suitable for you!" or" No information found" as shown in figure 1.

Fig. 1. Screen-shots of Preferences Monitoring system

Users also can write comments on their profiles regarding a certain product, and view other users' comments. They can also receive notifications about similar products to the one they usually buy or scan its barcode. Furthermore, users are able to see history of last mobile recognition. Our system was initially developed for nonnative consumer who cannot read Chinese in Taiwan, and then it included health monitoring services. The first part of the system has completed, while the second part is still under testing and will be ready at the end of May, 2011.

4 System Architecture

The general architecture of our system is shown in Figure 2. The system contains three parts: the user profile side, the mobile phone side - barcode recognition

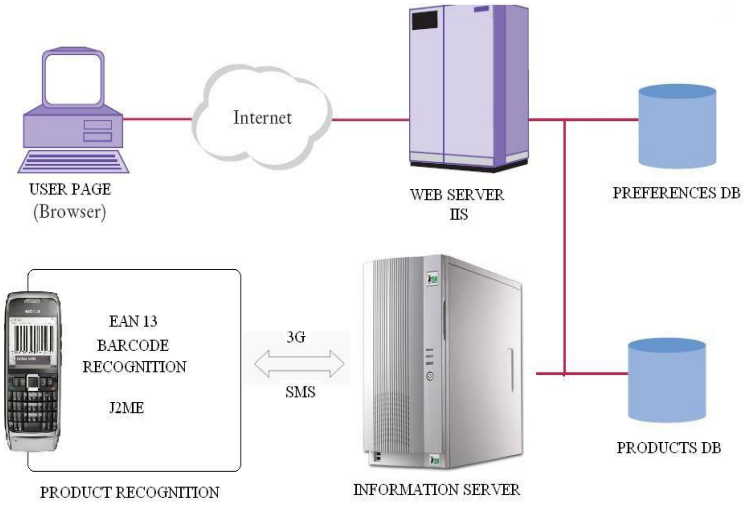


Fig. 2. System Architecture

component running entirely on J2ME enabled-mobile phone that supports the MMAPI9 (Mobile Media APIextension) extension, and the server side including web server and information server component which is located on a separate server.

4.1 User Profile Side

The user pages are designed by ASP.NET and communicate with web server and database by the internet. The pages allow users to predefine their profile, enters user food culture, e.g. Muslims can eat only Halal food, Hindus are vegetarian, pork and shellfish are strictly forbidden in Judaism, etc, the substances they are allergic to, calories based diet, consumer likes and dislikes and any details the consumers want. Users also can view and define other services on their accounts.

4.2 Mobile-Phone Side

The mobile phone enabled barcode recognition offers functionalities to recognize an EAN-13 barcode, communicate with servers and display results. The mobile phone side contains predefined J2ME components (libraries) for the major tasks: Object recognition, communication with a remote server as well as an XML parser for presenting the final information to the user using the available graphical user interface. The algorithm which is responsible for barcode recognition is based on the Batoo-Toolkit [11]. It is designed for Nokia phones. Additional components (functional modules) can be easily added without any required knowledge about the image recognition task. For example, the communication between the mobile phone and servers are based mainly on the 3G, other communication types such as SMS can be easily added.

4.3 Server Side

The server side contains the web server that is responsible for communicating with the web server and providing the user profile side with the required information and functionalities. The second part is the information server which is responsible for communicating with mobile phones and product database. It receives requests from users and matches the request with user profile and product information. Then, it returns the result to the mobile phone as an XML file when the communication is 3G. Sending an XML file instead of bytes reduces the communication delay time.

5 Prototype Testing

We have finished the first part of testing. We tested the system on a simulated database of markets with foreign students living in Taiwan. International students in Taiwan are from diverse cultures and they have different preferences. The students used their Nokia mobile phones to scan barcodes of different products, send their request to a server using 3G communication and get immediate results.

The preliminary testing results are satisfying since the student showed a considerable willing to use the system because it eases their lives in Taiwan. Results also show that the process of barcode recognition was easy. The quality of image-processed barcode was enough to reduce the error rate of scanning and matching the product in the database. The communication speed was high with minimum delay. One limitation deserves mentioning here is battery consumption. We notice that some mobile phones consume more energy during the process of scanning barcodes and that is because of allocating more power to the camera.

The second part of testing will focus on integrating all parts of the system. Testing will contain real database of some supermarkets after making agreement with them. It will also contain testing health preferences with local people from Taiwan.

6 Conclusions

In this paper, we present a system that eases consumer shopping and maintains consumer preferences. Because people recently have been willing to see more services that solve their daily problems on their mobile phones, we propose “Preferences Monitoring” system. The system enables foreign and local people in a country to maintain their preferences, cultures, religions and health routines while shopping. After testing the system, we conclude that integrating the used barcode system in markets with modern mobile phones and middleware technologies has the potential for improving people lives. However, there are still issues in this field requiring more consideration such as supporting different mobile phone platforms, implementing efficient barcode recognition algorithm, and deploying up-to-date middleware technologies. Our future work will contain testing more services of the system in real databases and integrating the system with multiple mobile phone platforms such as Android and Windows.

References

1. International Organization for Standardization: Information Technology International Symbology Specication-Data Matrix. ISO/IECI6022 (2000)
2. Junaini, S., Abduallah, J.: MyMobiHalal 2.0: Malaysian Mobile Halal Product Verification using Camera Phone Barcode Scanning and MMS. In: Proceedings of the International Conference on Computer and Communication Engineering (2008)
3. Damper, R.I., Garner, D., Jordan, G., Rahman, A., Saunders, C.: A Barcode-Scannaeird for Visually-Impaired People. In: 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Amsterdam (1996)
4. Liu, X., Lin, C.: Information Management System of Grocery Production Processing Based on a Bar Code Identification Technology. In: IEEE International Workshop on Anti-Counterfeiting, Security, Identification, pp. 164–168 (2007)
5. Ohbuchi, E., Hanaizumi, H., Hock, L.A.: Barcode Readers Using the Camera Device in Mobile Phones. In: International Conference on Cyberworlds, pp. 260–265 (2004)
6. Liang, Y.-H., Wang, Z.-Y., Cao, X.-Y., Xu, X.-W.: Real Time Recognition of 2D Bar Codes in Complex Image Conditions. In: International Conference on Machine Learning and Cybernetics, vol. 3, pp. 1699–1704 (2007)
7. Adelmann, R.: Mobile Phone Based Interaction with Everyday Products - On the Go, Next Generation Mobile Applications. In: The 2007 International Conference on Services and Technologies, NGMAST 2007, pp. 63–69 (2007)
8. Falas, T., Kashani, H.: Two-Dimensional Bar-Code Decoding with Camera-Equipped Mobile Phones. In: Fifth Annual IEEE International Conference on Pervasive Computing and Communications, pp. 597–600 (2007)
9. Sun, A., Sun, Y., Liu, C.: The QR-code Reorganization in Illegible Snapshots Taken by Mobile Phones. In: Gervasi, O., Gavrilova, M.L. (eds.) ICCSA 2007, Part III. LNCS, vol. 4707, pp. 532–538. Springer, Heidelberg (2007)
10. Awwad, Y., Lin, H.T., Yuan, S.M.: Hybrid Teaching Assistant for Developing Students Interactivity and Comprehension. In: Ninth IEEE International Symposium on Multimedia Workshops, ISMW 2007 (2007)
11. Adelman, R., Langheinric, M., Floerkemeier, C.: A toolkit for Barcode Recognition and Resolving on Camera Phones - Jump Starting the Internet of Things. In: Workshop Mobile and Embedded Interactive Systems (MEIS 2006) at Informatik (2006)

Practice and Study of Synergetic Learning Theory in Java Language Teaching

Zhihui Wang, Yuanqiang Wang, Shangfu Hao, and Baili Sun

Hebei North University,
Hebei, Zhangjiakou 075000, P.R. China
{iowzh,hbnuwyq}@126.com,
zjkcxh5656@sina.com,
skutu@163.com

Abstract. This article deals with problems of defective teaching effects widely existing in Chinese collage and university Java language teaching courses, and proposes a brand new teaching methodology by introducing two emerging learning patterns, i.e. synergetic learning and task-driven learning, into the classroom teaching. By conducting contrastive teaching experiments and analyzing teaching performance, the feasibility and maneuverability of this teaching methodology has been examined and demonstrated.

Keywords: synergetic learning; task-driven learning; Java language; teaching effect.

1 Traditional Teaching Methodology and Present Problems

Because of the safe stable and portable performance of Java language and the huge demand for IT developers, a lot of colleges and universities have included Java language course on teaching plan, as one of those compulsory courses of computer science. However, in teaching practice, influenced by traditional conception of teacher-centrism, text-centrism and classroom-centrism[1], the principal role of learners are often ignored, causing undesired consequence of lacking integrity and coherence in the course of learners mastering and applying the knowledge. This problem is intensively reflected in the fact that learners could easily pass the examination and even gain very high score after learning the course, but when facing a little more complicated Java programs, they just don't know how to begin to solve these problems, not to mention conducting practical software development independently. This situation apparently fails to achieve the teaching objectives. How to make the most of the learners' principal role in the learning procedure to enhance their ability of developing software by applying Java language? This article explores a new teaching methodology of improving learning efficiency by introducing two emerging learning patterns i.e. synergetic learning and task-driven learning into Java language classroom teaching practice.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 303–307, 2011.

2 Connotation and Characteristics of Synergetic Learning and Task-Driven Learning

Synergetic Learning is the outcome of applying Synergetic Theory to the field of knowledge learning, and is a learning technology system framework [1] aiming to the knowledge era. That is, for example, under certain circumstance of learning, learners could build up a learning community for interacting and complementing with each other to improve overall learning performance. In synergetic learning, the knowledge structure of a learner could be shared with the entire group[3], compared with individual learning, synergetic learning could effectively improve the learning efficiency of the learners[4]. Task-driven learning is a teaching strategy based on constructivism learning theory[5], and it converts teaching-oriented strategy into task-oriented learning strategy. During the teaching practice, by assigning interesting specific tasks, educator motivated the learners to proactively understand the core of the knowledge as fulfilling their tasks.

Essentially, synergetic learning and task-driven learning are both of autonomous learning. There is a common purpose of these two learning patterns, i.e. to make the most of learners' principal roles while avoid disadvantages of autonomous learning, and the only difference between them lies in the specified ways of realization By establishing a mechanism of "coordination" "complementation" and "cooperation", the synergetic learning could strengthen the contact insider the learning community, and enhance the "centripetal" and "cohesion", so as to achieve a learning effect of "individual driven by and progressing with the collective". Picturesquely speaking, synergetic learning is the strategy of promoting learning through "internal force"; While task-driven learning is the strategy that through stimulating learners' learning interest and enthusiasm, educators could facilitate the learners active learning, independent thinking and achieve better learning effects, i.e. driven by "external force". By comparing these two learning patterns, it is not hard to discover that some disadvantages still remain besides their advantages. On the one hand, the community under synergetic model is completely in the specific condition of self-discipline and loose coupling, and usually results in higher failure rate esp. in the learning communities composed of on-campus students in our country, even succeeded communities are not repeatable; on the other hand, task-driven learning, in learner's point of view, is a strategy of "learn to use", and therefore will inevitably lead to one-sidedness of learner's knowledge structure.

In conclusion, for the sake of common function and advantage complementation of "internal force" and "external force", these two learning patterns share naturally overlapping portion. The author uses bold way of combining there two learning patterns in classroom teaching environment and receives good teaching effect.

3 Application Practice in Classroom Teaching

Synergetic learning and task-driven learning are both emerging learning patterns in the field of education, ways of how to employ them in classroom teaching environment effectively are still been explored. For this reason, in classroom teaching practice, we carried out teaching experiments to figure out a set of complete, feasible

new teaching method by contrasting with traditional teaching patterns. specific steps and contents of the teaching experimentation that carried out by the author are listed as follows:

3.1 Identify Specific Assignment According to the Training Objective of the Course

Java is a very practical computer programming language, and the training objective of Java language course, is to enable learners to develop system application software of certain practical value independently or cooperatively. For this reason, it can be more appropriate to choose a concrete real software development project as the "task", because of its practical and incentive value. In this teaching experiment, we chose two software programming projects of "Teaching & Research Service Platform for Middle Schools in Xuanhua District" and "Video Synergetic Learning Platform for Computer Organization and Architectures" as the learning tasks.

3.2 Identify Periodical Objectives and Figure Out the Learning Sequence by Combining with the Demanded Knowledge for the Task

Software development project is a complicated systematic engineering, and according to software engineering basic theory, full analysis should be carried out based on project business process, to determine the basic function module and frame structure, and formulate development schedule in accordance with the project time requirement. It must be pointed out that all participants involved in this teaching experiment are beginners without any developing experience, for this reason, teachers should summarize and analyze Java language knowledge points demanded for each function module while conducting demand analysis, and clearly mark the name of demanded knowledge point, location or means to obtain data as well. By this means, learners could figure out a clue of learning sequence out of the messy Java knowledge system, to achieve the ultimate learning goals more easily.

3.3 Establish Synergetic Learning Platform

Comparing with the traditional narrative teaching mode, synergetic learning could make the most of learners' principal roles to achieve coordinated paces, and a synergetic learning platform is indispensable. We established the synergetic learning platform by taking following measures:

1) Share podium, blackboard and chalks, which used to be teacher's privilege, with the learners, then transfer the classroom into learners' forum and switch teacher's role to activity organizer.
2) Establish Java course BBS site on campus LAN, co-managed by teachers and learners.
3) Referring to Johnson brothers' research achievement of "group cooperative learning", to divide learners of a grade into groups consisted of 4 to 5 persons according to the principle of character complementary.
4) Inside the group, there will be an interactive activity about Java every two days, to discuss about what they'd learnt about Java programming skills

during the period or what they came across in programming. After the activity, group recorder should issue the minutes onto the campus Java course BBS as a "post".

5) Teachers would sort and screen the posts on the course BBS aperiodically, to collect questions and problems that learners run into during the course, and organize further discussion over the above issues in the class, for those left unsolved or undetermined, teachers should not give the answer immediately, but to give them hints on where to get sufficient materials and how to draw conclusion. Final results could only be obtained by learners conducting experiment or retrieving data.

3.4 Establish Task Supervision Mechanism and Periodical Incentive Mechanism in Accordance with the Task Schedule

According to the project schedule, allocate specific periodical assignment to every student participating the experiment, and underline the deadline for the task. During the course, by periodically enquiring the learners or checking the code, teachers should promptly recognize problems and difficulties that the learners come across to supervise and promote the progress; teachers should promptly comment the performance to motivate learners for further study every time a periodical assignment was accomplished.

4 Analysis and Conclusion on Teaching Experiment Result

The purpose of Java programming course is to arm the learners with preliminary Java software programming ability. And the programming ability reflects a learner's comprehensive ability, for this reason we shall state our point of view from three aspects as listed below

4.1 Contrast on Learners' Direct Advancement

By conducting this teaching experiment, we have discovered that five abilities of the participants have apparently improvement as shown in Table 1.

4.2 Teaching Achievement Aspects

There are two teaching project programmed in this experiment, "Teaching & Research Service Platform for Middle Schools in Xuanhua District" has been officially put into use, and has won users' praise; "Video Synergetic Learning Platform for Computer Organization and Architectures" was brought to The National Tenth Multimedia Courseware Competition and won the second prize.

4.3 Employment Status of Graduates

By the time this article was complete, the first batch of the learners participating the teaching experiment had already graduated, all of which are working as Java

developers in Beijing except one in institutional organization; the second batch is just about to graduate, yet all of them are employed by some famous corporation dedicated in the field of e-book management software development.

Base on the analysis of the three aspects listed above, the author believes that the classroom teaching experiment of introducing synergetic learning and task-driven learning patterns into Java programming course is comparatively successful. However, it must pointed out that only a minority of students participated in the experiment, and a great many work must be done to formally apply this into classroom teaching environment.

Table 1. Comparison of teaching experiment performance

Items	Experimental group		Traditional group	
	Before experiment	After experiment	Before the course	After the course
Information retrieval capability	Below average	*Good*	Below average	Below average
Communication ability	Average	*Above average*	Average	Average
Coordination and cooperation skills	Average	*Above average*	Average	Average
Analytical skills	average	*Good*	Average	Average
Program execution efficiency	Below average	*Above average*	Below average	Average

References

1. Zhu, Z.: Synergetic Learning: Learning Technology System Framework Aiming to the Knowledge Era. J. Chinese Audio-Visual Education 4, 5–9 (2006)
2. Jiang, W.: A Theoretical and Practical Research on Enhancing Classroom Synergetic Learning. J. Education Theory and Practice 8, 49–53 (2006)
3. Chen, Q., Liu, R., Zhang, J.: Creating a constructive environment for learners to explore word problem solving. In: International Conference on Computers in Education, Beijing (1998)
4. Liu, J., et al.: An Generating Method of Synergetic Learning Task for Individual Network Learning. Journal of Xi'an Jiaotong University 8, 801–803 (2004)
5. Wang, Z.: A Brief Exploration and Practical Research on Task-driven University Teachers Education Technology Training. J. Chinese Audio-Visual Education 12, 104–105 (2009)

Personalized Graphical User Interface Based on Java

Ning Xiao, Xiaojun Wang, and Yaming Han

Shanxi University of Finance & Economics,
No.108, Jinyang Street, Taiyuan, Shanxi, China
abcdxn@126.com

Abstract. Java provides various tools about graphical user interface, but its function is weaker than other languages of Microsoft. A simple and effective method was provided by Java. Through rewrite Windows form and its behavior, it can generate custom personalized graphical user interface. This method can be applied to development of various interactive software user interfaces. Two kind of personalized graphical user interface effects have been realized in this paper to satisfy needs of most personalized graphical user interfaces.

Keywords: Graphical user interface; Personalized; Java; Windows form.

Along with the rapid development of computer technology, computer application has been wide spread. The graphical user interface has greatly improved the efficiency of non-professional user. The graphical elements of graphical user interface and the relations between them is a metaphor of information features in the real world [1][2]. Users can have interaction with interface by direct operation. Java provides various tools about graphical user interface, but its function is weaker than other languages of Microsoft such as VB and VC++.

1 Standard Windows Form

Form is visualization graphic for displaying information to user in the Windows applications. A formalized standard Windows form consists of three parts. The top of form is title bar. The second part is the main part of the form, and the term is the client area of the form. The remaining surrounding the main periphery of the form is form border.

The form's events are mainly embodied in the title bar and form border.

(1) Title bar basically has several events as following:

i. Maximization, minimize events. These are mainly used to control the size of the form. It can switch between maximum and minimum state.

ii. The shutdown event. The form will be closed when click close button.

iii. Drag events. The form will move when press left-click on the title bar and drag mouse.

(2) Border events, are primarily used to control the size of the form. Keeping left-click when the cursor arrive the range of form's border can change the size of form continuously.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 308–313, 2011.

2 Rewrite Windows Form, Create Personalized Form

Form represents container class, because the container's appearance can be only influenced by form in practical applications. Our own personalized form can be created.

2.1 Rewrite Title Bar

Java has its own setting about Frame, there is detailed narrative in [3].We use rewriting to realize the same function as below.

Title bar can inherit Panel, and then set its default length and width. The construct function of title bar should accept a parameter of JFrame Class. So we can make use of some original methods which belong to JFrame Class, and then get length and width by the parameters to set length and width of the title.

Then rewrite title bar icon. SetImageIcon of JFrame or Frame can be used to change JAVA default icon and JLabel components can be used to replace containers of the icon, because JLabel can not only set text and pictures but also set whether transparent or not by the opupe method. This Jlabel will be added to the Jpanel namely title bar, and can be rewrote with paint method to add some lovely background pictures. But the paint method cover design and text of the Jlabel, we can not use paint method directly to paint background picture. In order to solve the problem, we can create a picture zone in memory by BufferImage, and then get Graphics from the picture area. First, draw the background images needed, and then draw BufferImage by getting Graphics from Jpanel. So it will not happen that the background picture of Jpanel covers the Logo icon of JPanel. Finally rewrite the SetImageIcon method of JFrame to set the Logo icon of JLabel, while these methods usage are consistence with former method.

Three common buttons of the title bar should be rewritten. We choose JButton to realize button of the title. JButton has powerful functions, and can be used to create personalized button.

The main code of title bar is as follows:

```
Class BiTiLan extends JPanel {
  JButton DXG[] = new JButton[3];
  // maximization, minimize and close button
  JLabel Logo = new JLabel();
  JFrame parent;
  BiTiLan( JFrame parent ) {
    // the parent is the attachment form of BiTiLan
    this.parent = parent;
    …      // settings related DXG
    …      // settings related Logo
    …      // settings related this
  }
  Public void paint( Graphics g ){
    BufferImage b =
Buffer.CreatImage(0,0,Width,Height);
    // in memory create double buffering BufferImage
    Graphics gg = b.getGraphics( );
```

```
    gg.drawImage(  /* own figure */  );
    g.drawImage(  /* draw BufferImage*/  );
    /* realize the whole could avoid the image in the
JPanel covering the pictures of JLabel */
  }
}
```

Whole title bar's rewrite job has been finished.

2.2 Take the Title Bar into Our Own Form

i. First, take off the title bar of JFrame, the method is setUndecorated(true).
ii. Set Layout of our own form null, this step is important.
iii. Take title bar into the JFrame which has rewritten.
Next step we have to rewrite the behavior.

2.3 Rewrite the Behavior of the Title Bar

Title bar has two kinds of behavior: Press left and drag the mouse in the title bar, then the form will be moved, drag form is short for it; Double click in the title bar, form will be maximized state.

To realize the behavior of dragging form, the simplest behavior is adding the title bar with a mouse listener. Set a Boolean switch variable which it opens when left click. Now when the mouse is moved to reach a given position, relax the left key, and set Boolean switch closed, then resetting form's position base on the mouse's position. In this case, the largest malpractice is that the form cannot move smooth and has a certain jumping.

A better method is that adding a thread to the program. Let this thread repeat the above process constantly during dragging, so the form can move smooth.

Actually, there is a better kind of method can also get fluent effect, which is calling API which Windows obligate. But we do this in order to preserve platform irrelevance which Java owned.

The behavior of realizing double-click maximize is simple. Add the mouse listener to the title bar, and judge, if the combo time > 2, then set the window in maximization.

A further study of its behavior find that, according to the above method of realization the title bar's size is not be changed synchronously when the form's size is changed. To solve this problem, we should add container listener to the title bar. The container listeners are made for listening containers, it will make some corresponding detail processing when container's size is changed. Then we can add program to the corresponding method to make the pictures' size in the title bar changed corresponding.

2.4 Form Border

It is the edge of around that plays an important role for a form in addition to the most important title bar. These edges have two purposes: It is to make the pictures more perfect and full in the perspective of computer graphics; they have a particularly important function which can change the size of the form continuously.

There are two methods for realizing edge:

i. Create edge in the BorderFactor and add it to the form, then create a thread to judge the position of the mouse constantly. If its horizontal y-coordinate reaches a certain condition, the change of the mouse's shape gives users the tip that they can change the size of the form.

ii. Add several small JPanel to make edges, and add mouse listeners to them respectively. When the mouse comes in, the change of the mouse's shape gives users the tip that they can change the size of the form.

2.5 Realize the Personalized Window

Methods related to the form can be rewritten after finishing the work above, then package in jar file. So far, we have completed the personalized design and production for the form, it can maximum disassembling and restructuring in accordance with the needs of the individuation.

Figure 1 is an example of personalized window.

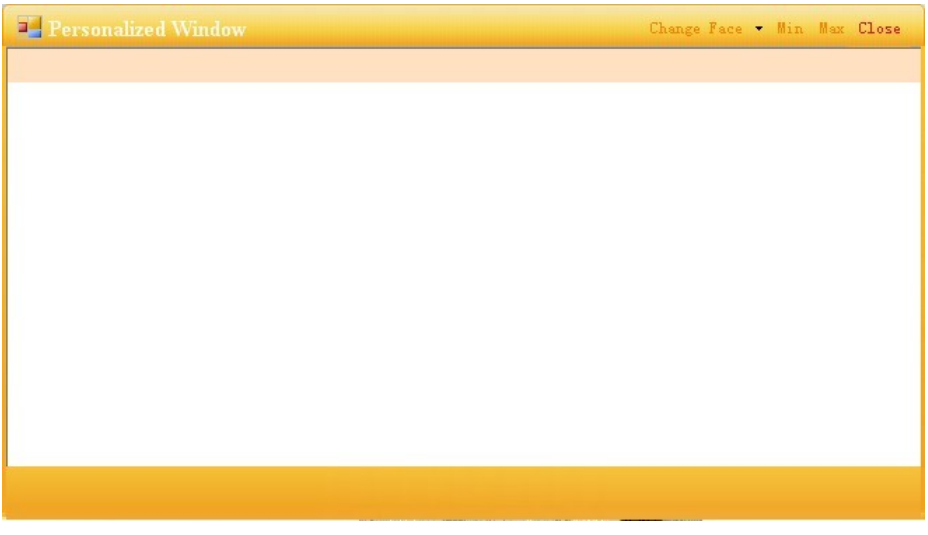


Fig. 1. An example of personalized window

3 The Realization of Non-standard Windows Form

The realization process of irregular form in Java is discussed in this section.

As we known, the API for setting form graphics irregular is AWTVtilities.SeWindowShape(window, shape), so the key step to create irregular form is how to create an irregular shape.

Standard shape such as elliptic and rectangle can be created by corresponding method provided by Graphics2D.

If Graphics2D cannot draw, we can use pictures to draw form effect needed, and then turn the pictures into the shape graphics which are needed.

Steps for realizing: get picture; get information of the pictures by “get pixels”; get every pixel transparency, if transparency is 0, then abandon the pixels, if transparency is 1, then get the pixel; use these data to create a shape graphics; set window’s shape.

According to this method a personalized form instance is shown in figure 2.

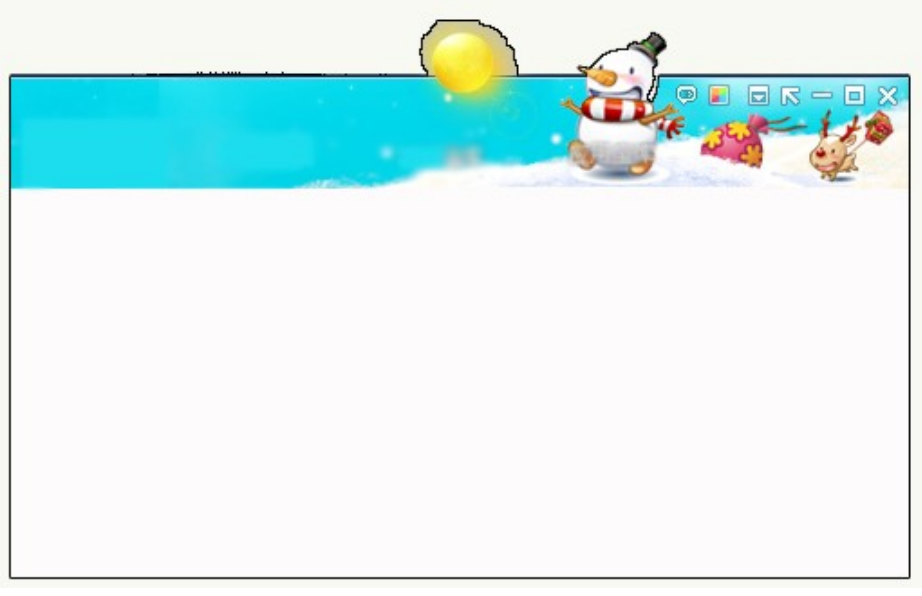


Fig. 2. Non-standard Windows form realization

The core codes of personalized examples form in Figure 2 are as below:

```
public Shape getImageShape(Image img) {
  ArrayList<Integer> x = new ArrayList<Integer>();
  ArrayList<Integer> y = new ArrayList<Integer>();
  int width = img.getWidth(null);// image width
  int height = img.getHeight(null);// image heigh
  // screening pixel
  // acquire all information of the image’s pixel
  PixelGrabber pgr = new PixelGrabber(img, 0, 0, -1, -
1, true);
  try {
    pgr.grabPixels();
  } catch (InterruptedException ex) {
    ex.getStackTrace();
  }
  int[] pixels = (int[]) pgr.getPixels();

  // circulation pixel
```

```
  for (int i = 0; i < pixels.length; i++) {
  /* screening, put the opaque pixel's coordinates into
coordinates of ArrayList x and y */
    …
  }

  /* Establish image matrix and initialization (0 is
transparent, 1 is opaque) */

  int[][] matrix = new int[height][width];
  for (int i = 0; i < height; i++) {
    …
  }

  /* import the opaque coordinate information of
ArrayList */
  for (int c = 0; c < x.size(); c++) {
    matrix[y.get(c)][x.get(c)] = 1;
  }

  /* Because the area class indicated by the Area can
be combined, "scan" each line of the image rectangle
horizontally, turn opaque pixel into Rectangle, then
combine Rectangle in each line by rec's object of Area,
finally form a complete Shape graphics. */
  Area rec = new Area();
  int temp = 0;

  for (int i = 0; i < height; i++) {
    for (int j = 0; j < width; j++) {
      …
    }
    return rec;
  }
```

4 Conclusion

Rewrite Windows form through the use of Java can realize various personalized Windows form. The two different effects realized by this paper can satisfy needs for most personalized graphical user interface.

References

1. Dix, A.: Human-computer Interaction. Addison-Wesley, Reading (1998)
2. Liu, Q.-f.: Java-based Object-oriented Interactive Graphics Toolkit. Computer Engineering 35(15), 250–252 (2009)
3. Loy, M., Eckstein, R., Wood, D., Elliott, J., Cole, B.: Java Swing. Oreilly & Associates Inc., Sebastopol (2004)

Integrated Teaching Platform of Computer Control System Based on Matlab/RTW*

Shaoqiang Yuan, Huihui Wang, and Gongzhang Shen

School of Automatic Science and Electrical Engineering,
Beihang University, Beijing,100191, China
ysq1960@sina.com

Abstract. To improve the quality of teaching experiment for computer control system, an integrated teaching platform based on Matlab/RTW is given in this paper. It can easily translate Simulink blocks to C/C++ code easily. The process of system design, simulation and experiment can be conducted smoothly in one integration platform. This paper describes the semi-physical simulation in detail, and demonstrates the application of the platform in teaching practice. The Teaching Experimental Center of Automatic Control in BeiHang University has successfully applied the integrated teaching platform to experiment teaching. Through teaching practice, the teaching platform shows some advantages, including high efficiency, low cost, friendly interface and the function of testing dynamic process and rework parameters on line. Therefore, it plays a positive role in improving the teaching level and quality.

Keywords: integrated teaching platform; RTW; semi-physical simulation; real-time control system.

1 Introduction

Experimental teaching is an important part of computer control teaching. It is advisable for students to follow the instructions as given below in the traditional experimental teaching. First, students should build a mathematical model for control system. Then, students are required to implement off-line simulation in Simulink environment to develop the appropriate control law [1]. Finally, students must write the real-time programs according to the control law, and run programs on the hardware. It is verified in practice that this method has several serious drawbacks: (1) Off-line simulation is a kind of simulation with simple mathematic model. Simulation environment turns out to be quite different from physical environment, so traditional design stage and implementation phase have been disconnected. (2) Without the parameters-editing function, the written programs can't execute real-time monitoring. The labs have to use expensive monitoring equipment while debugging, which leads to a waste of time and money. (3) With lots of difficulties in designing, it always takes a long time for students to accomplish the real-time control program. (4) Since the experiment methods are dull and difficult, students have no interest in the experiment. Over time, the effect of teaching can't be guaranteed.

* This work is supported by Construction Projects of Beijing Excellent Teaching Team.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 314–320, 2011.

This paper presents a new integrated teaching platform of computer control system, which can accomplish the whole design process rapidly by a powerful tool of Math Works. The process of system design, simulation and experiment could be conducted smoothly in one integration platform. Simulink model is regarded as interface of connection physical objective; it is compiled to the executive code with Real Time Workshop (RTW); the simulation results can be ported to actual system directly to improve further. So the platform has great significance for further improving the experimental teaching, raising experimental level and raising working efficiency.

2 The Design of Integrated Teaching Platform

2.1 The Design and Development Flows of Integrated Teaching Platform

Figure 1 shows the process of traditional experimental method, while Figure 2 illustrates the new experimental method based on integrated teaching platform (the gray segments are operated in Matlab environment). As shown in Fig. 1, the new experimental method is composed of the following steps [1], [2]:

1) Establish digital simulation models based on system design. Instead of building complicated mathematic model, the students can build mechanical models using SimMechanics toolbox of Matlab.

2) During the period of digital simulation, students can try different kinds of control laws, and determine the control scheme initially. It must be added that intelligent control can be realized in this stage by using Matlab toolboxes.

3) Semi-physical simulation is the most important part in the design and development flows. The Simulink model need to be changed to executable code [3] by RTW, then the C code can be downloaded into target machine. Then students should replace the mathematics mode for actual control object. The so-called semi-physical simulation is that the actual circuit boards are embedded in the simulation loop, so virtual controller and physical environment are connected together. The parameters can be modified online and the control law can be optimized during this stage.

4) At last, all the control system is required to be divorced from Matlab and be realized by practical circuits.

As can be seen from the above steps, the teaching platform emphasizes on integration of control system design. The so-called integrated design is to integrate control system modelling, digital simulation, and semi-physical simulation in a platform.

2.2 The Composition of the Integrated Teaching Platform

Integrated teaching platform consists of two parts: one is software environment comprised by computers and related software, the other is factual environment composed of controlled objects and other peripheral circuits. The two parts are connected by I/O drive module.

To realize the integration of simulation, it is critical to select the proper software. Matlab, which is widely applied in control field, is chosen as software platform in this paper. Simulink, a visual graphics software package of Matlab, provides a graphical interface for dynamic system modeling, simulation and analysis. RTW, an important

supplement function module for MATLAB and Simulink, provides a real-time development environment [4]. It can automatically translate modularized models to codes, run models of dynamic systems on the hardware, and debug programs based on models. Real-Time Windows Target, an additional product of RTW, uses a small real-time kernel to ensure the realization of real-time operations. It makes Simulink and the generated codes run in the same computer.

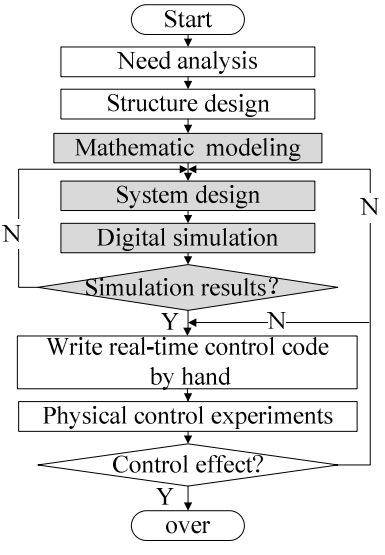


Fig. 1. The flow of traditional control system experiment

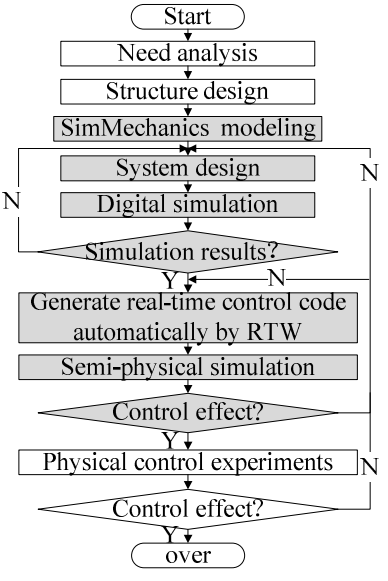


Fig. 2. The flow of teaching method based on the integrated teaching platform

To realize the integration of teaching platform, it's essential to choose the hardware interface board and development of drivers. Real-Time Windows Target provides a wealth of driver models, which can support hundreds of I/O boards. Students can develop their own drivers by the S-function models of Simulink for the card which RTW doesn't support.

2.3 The Scheme and Characteristics of Semi-physical Simulation

There are two schemes during semi-physical simulation stage: one is that the object programs from simulation computer are downloaded to external target computer; the other is that simulation computer and target computer is the same computer. The latter scheme is taken in this teaching platform. This scheme has the following features [5]:

1) Compared to offline simulation, conditions of semi-physical simulation are closer to the actual conditions. Hence, semi-physical simulation is more conducive for control system design.

2) Real-time control codes can be generated directly from the Simulink graphical models. It is helpful to eliminate students' tedious coding process.

3) Simulation computer and target computer is the same computer. It saves funds and reduces download time.

4) It is simple and intuitive that Simulink screen is used as the panel of real-time monitoring and debugging.

5) With the function of modifying real-time parameters online, students can test and validate their designs in real-time environment while experimenting.

2.4 Realization of the Semi-physical Simulation

Semi-physical simulation system is established from offline simulation by some steps as follows [1]: abolishing the mathematics model of the object, adding the interface module of Simulink I/O to achieve communication with hardware by linking to data acquisition card, linking hardware to simulation loop. Site signal will give a feedback to workspace in control system; control variable outputs with Matlab will transmit to the executive faculty.

In the semi-physical simulation, it not only requires installing the Real-Time Windows Target kernel and selecting the compiler, but also needs to set parameters such as simulation time, step and integrators. Both codes' automatic generation and dynamic process monitoring, two important parts of semi-physical simulation, are needed to set related parameters in the Matlab environment [5].

The exact way of code generation is shown as follows. Only when the designed results meet the requirements very well during off-line simulation stage can student remove the controlled object modules, connect the corresponding interface modules, set simulation mode to external mode, and change the target file system to rtwin.tlc. After that, students can double-click the build icon to generate real-time control program [6].

The detail of dynamic process monitoring is shown as follows. After external mode is selected, students can double-click the Start Real-Time Code icon to start the real-time simulation. To ensure that the system can be monitored, oscilloscopes should be opened up. And when the parameters of the controller modules (PID module, for example) are modified, parameters in the control system are changed, leading to the variation of the output signal.

3 Application of the Integrated Platform in Teaching Practice

Automatic Control Theory Laboratory, part of Teaching Experiment Center of Automatic Control in BeiHang University, has applied the technology of integrated teaching platform to experiment teaching. Using the integrated teaching platform, the author develops a number of simulation control objects, such as inverted Pendulum, bent-arm Pendubot, bridge crane and so on. These developments can be used for experimental teaching directly. Owing to space constraints, this paper only shows two examples of control system to illustrate the integrated platform.

3.1 Bridge Crane Experiment

Bridge crane control system developed by BeiHang university is the first computer-controlled crane model in China [7]. The working principle is as follows. DC motors drive the car to move on the track horizontally by transmission system. There is a

hanging weight below the car. The pendulum angle of weight and the car's position can be measured by the potentiometer, while the car's speed can be measured by tachometer. This system requires that the pendulum angle stops swinging as quickly as possible after the car's movement. The linear mathematical model can be described as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\dfrac{K_e K_t}{MR_a r^2 + R_a J} & -\dfrac{mgr^2}{Mr^2 + J} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\dfrac{K_e K_t}{\mathrm{l}(MR_a r^2 + R_a J)} & -\dfrac{(M+m)gr^2 + gJ}{l(Mr^2 + J)} & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{rK_t}{R_a(Mr^2 + J)} \\ 0 \\ \dfrac{rK_t}{lR_a(Mr^2 + J)} \end{bmatrix} .U_a \qquad (1)$$

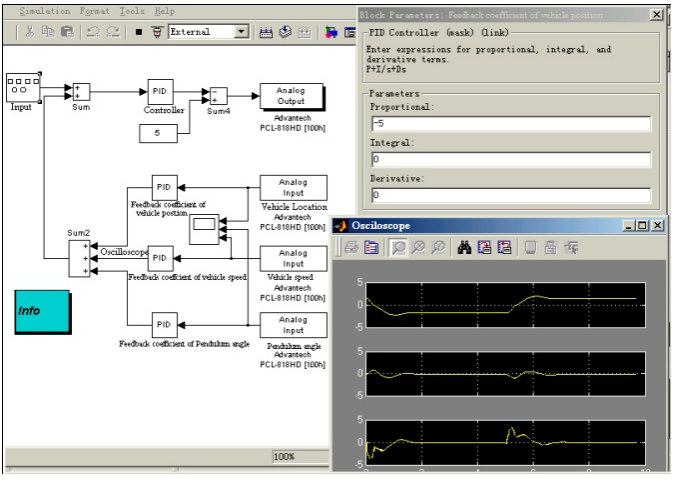


Fig. 3. Real-time controlling, monitoring and online parameter adjusting interface of RTW

In the experiment, sampling time is 0.02 seconds, while simulation time is 10 seconds. PCL-818HD of Advantech Corporation is used as analog input and output board. Students should follow the process of integrated teaching platform mentioned above. During the semi-physical simulation, control output variables with Matlab are transmitted to the executive faculty, while site signals give a feedback to workspace. To monitor the real-time dynamic process and modify parameters on line, students need to set relevant parameters, as shown in Fig. 3. Finally, the bridge crane experiment gains the desired control effect in the integrated teaching platform.

Bridge crane control system has been used as teaching experiment system for a long time in the Teaching Experimental Center of Automatic Control. It's more effective to apply the integrated platform to the teaching of crane control system. Now, it has been widely used in many fields, such as experiments teaching, curricula design and graduation design.

3.2 Mixed Simulation Experiment Based on Teaching Platform

Mixed simulation system is mainly composed of three parts: digital machine, analog machine and digital-analog conversion equipment [8]. Digital machine, which controls the whole operation of system, has the function of analyzing and judging the simulation results; Analog machine, including kinds of circuit devices, is a kind of integrated circuit board which is used to build a variety of control object. Digital-analog conversion equipment is another important part, and it is used for exchanging data between the previous two.

An experimental model of a second-order system is shown in Fig.4. Students should establish a Simulink model in the digital machine and build an analog circuit of second-order system in the analog machine. Then Normal mode is selected to implement theory simulation, while the External mode is chosen to perform experiments.

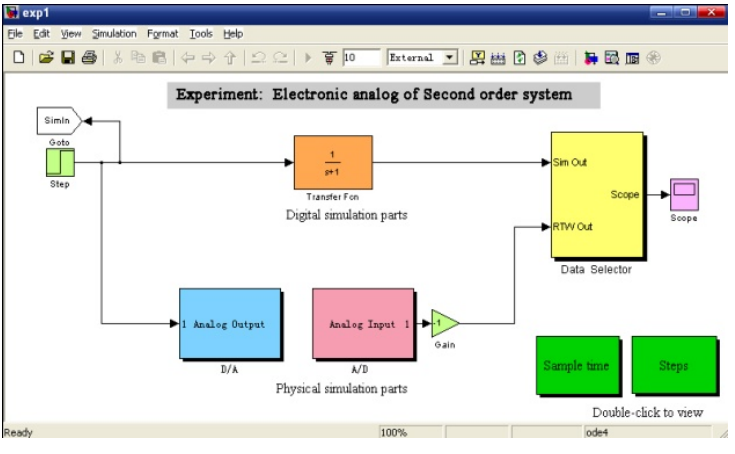


Fig. 4. Control system modeling interface

At present, more than two thousand students use the mixed simulation platform to complete the automatic control theory experiments at Automatic Control Center Laboratory of BeiHang University every year. And the mixed simulation system can complete all the exiting automatic control experiments, which reflects that integrated teaching platform is of universal applicability and flexibility.

4 Conclusion

This paper puts forth an integrated teaching platform based on Matlab/Simulink/RTW. The platform takes full advantage of Matlab. It can generate codes automatically by RTW and test dynamic process and rework parameters on line by the interface of Simulink. Practice shows that with the application of teaching platform, students can concentrate on the design and implementation of control system. The platform not only solves the traditional experimental problems, but also

has a good flexibility, autonomy and design ability, and increases the interests of students. Teaching Center of Automatic Control of BeiHang University has put this teaching platform into the teaching of computer control, and obtained good teaching results.

References

1. Song, Z., Wang, J., Liu, C., Song, X.: Design of a Hardware-in the Loop Experiment Simulation System for Process Control based on RTW. In: 2009 First International Workshop on Education Technology and Computer Science, pp. 1123–1126. IEEE Press, Wuhan (2009)
2. Li, G., Tang, Z.-Y., Geng, H.-X., Jin, F.-R.: New integrated design method for control system base on Matlab/RTW. China Measurement & Test 36(1), 75–77 (2010)
3. Luo, F., Huang, Z.: Embedded C Code Generation and Embedded Target Development Based on RTW-EC. Computer Science and Information Technology (ICCSIT) 5, 532 (2009)
4. Li, Y.-F., Yang, P., Zhang, Y.: Real-time Mixed Simulation Platform for Advanced Process Control Based on MATLABf. Journal of System Simulation 19(18), 4150–4153 (2007)
5. Yuan, S., Chang, P.: Design, Simulation, Experiment Integration Platform of Computer Control System Based on RTW. Journal of Electrical & Electronic Education 31, 65–68 (2006)
6. Real-Time workshop for user with SIMULINK. The Math Works, Inc. (2007)
7. Yuan, S., Chen, L., Li, X.: Study and application of semi-physical simulation condition based on RTW. Chinese Journal of Scientific Instrument 27(6), 2582–2586 (2007)
8. Yuan, S., Zhang, Z.-h., Chen, L., Li, X.-s., Li, B.-h.: Development & Application of Mixed Simulation System Based on RTW. Journal of System Simulation 21(3), 700–702 (2009)

Introducing Probability for Model-Based Cognitive Diagnosis of Students' Test Performance*

Junjie Xu and Rong Chen

College of Information Science and Technology,
Dalian Maritime University, Dalian 116026, China
{rchen.cs,tsmc.dmu}@gmail.com

Abstract. Cognitive diagnosis as an instance of model-based diagnosis, it works on the model of students' courses in terms of knowledge items that they may learn, tests them and helps them to understand their faults in cognition. In this paper, courses are formally defined as set of knowledge items with requirement constraints, knowledge items are associated with a set of *exam questions*. Moreover, diagnostic algorithms are used to help a student understand what knowledge item within a course the student does not master. Each knowledge item has a degree of correlation with knowledge items have not mastered, which can be used in computing suspicious degree of knowledge items as a selection criteria for the final results. Find the root reason of his/her test errors, and the recommendations like what should be done next. Experimental results show that the group of students with such understanding can improve their testing performance greatly in an E-learning environment.

Keywords: model-based cognitive diagnosis; course model; knowledge items; suspicious degree.

1 Introduction

A widely accepted definition of *cognitive diagnosis* has been given by Ohlsson [1]: "*cognitive diagnosis* is the process of inferring a person's cognitive state from his or her performance." Cognitive diagnostic modeling has become an exciting new field of psychometric research. Cognitive diagnostic models aim to diagnose students' status of a group of discretely defined skills, thereby providing them with detailed information regarding their specific strengths and weaknesses [2]. To date, at least fourty distinct cognitive diagnostic models have appeared in the literature [3].

This paper puts computerized cognitive diagnosis in a MBD framework; we recast an adaptive course testing as a diagnosis system in MBD; a course description is mapped onto the system description, knowledge items onto components, and observations are students' inputs that indicate their test performance. Here, courses are defined as set of knowledge items with requirement constraints, knowledge items

* This work is supported by National Natural Science Foundation of China (60775028), the Major Projects of Technology Bureau of Dalian No.2007A14GXD42, and IT Industry Development of Jilin Province.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 321–327, 2011.

are associated with set of exam questions, and thus diagnostic algorithms are used help the student improve learning outcomes. After we get the result of diagnostic algorithms, we can do a further processing through applying the way of probability and statistics to measure the degree of correlation about each two connected knowledge items. For diagnosis number more case, has the very vital significance.

The rest of the paper is organized as follows: Section 2 provides more discussion on related work. Section 3 presents our method on model-based cognitive diagnosis of students' test errors. In Section 4 we think about experiments and evaluation of our cognitive diagnostic system, and finally we conclude with some future work.

2 Related Work

2.1 Cognitive Diagnosis in Intelligent Tutoring Systems (ITS)

Cognitive diagnostic modeling has become an exciting new field of psychometric research. Cognitive diagnostic models can be used to detect the presence or absence of specific skills required for educational exams. In [3], authors show the performance of testing selection strategies based on Shannon entropy and Kullback-Leibler information are compared with random item selection. Dillenbourg proposes a framework for building a model in intelligent tutoring [4], which give a graphical representation of the learner model process. It is based upon the distinction between behavior, behavioral knowledge, and conceptual knowledge (in a 'vertical' dimension) and between the system, the learner, and the system's representation of the learner (in a 'horizontal' dimension). But this model just serves an analytic function rather than description, with a relative neglect of the relation of knowledge.

2.2 Diagnosis in AI

Sterbini and Temperini consider course configuration and propose a logical framework for describing courses in [5]. Their goal is not cognitive diagnosis, but to help a student to plan a set of courses which could make them learn the knowledge as quickly and completely as possible. In [6,7], a theorem prover is first used to compute conflict set, then the HS-tree algorithm runs upon such a set to compute diagnoses (i.e. hitting set). HS-tree is the original algorithm for computing diagnosis candidate, but it may lose the optimal solution by pruning. To overcome this, Greiner proposes a HS-DAG algorithm in [8]. Jiang et. al. also propose a BHS-tree algorithm, having the advantage of efficiency because of reduced node size, i.e., the node tree generated is significantly smaller than the HS-tree, and re-calculation is avoided in a new conflict set when the need completely re-calculated, only need to add the addition of new branches on the basis of original BHS-tree.

3 Model-Based Cognitive Diagnosis from Performance Evidence

3.1 Application MBCD to Computer Network Course

We introduce the architecture of course modeling, which has a three-tier architecture which is composed of the presentation layer (there is a GUI that users

can do exam question), the business layer (get questions from Test Data and diagnose) and the data access layer (store the Test Data and System Description). Diagnosis techniques are built upon the strict logical foundation provided by Reiter (in [6]). A course description is mapped onto the system description, knowledge items onto components, and observations are students' inputs that indicate their test performance.

We use an example to illustrate. The course Computer Network divided into 10 knowledge items as {MAC_addr, ip_addr, login, subnetting, route_selection, gateway, protocol/port, proxy_http/socks, proxy_connection, connection}, also named {C1,…, C10}. We describe the first layer model in Fig.1, each knowledge item is treated as a component, where each vertex is a component and every arc depicts the dependence between two components. For example, the arc from 'ip_addr' component to 'connection' component means that one of the required knowledge for 'connection' is 'ip_addr'. And all the relationship are shown in Table 1, in which "RK" is the required knowledges of Ci. E.g. "RK" of C1 is nil that means C1 has no required knowledges.

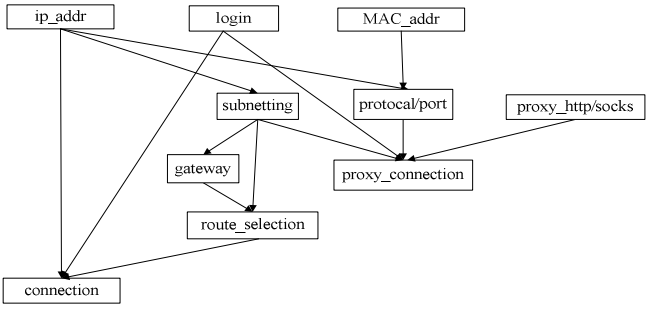


Fig. 2. A course model

Table 1. Learning components in a course mode

	Component name	Required Knowledges		Component name	Required Knowledges
C1	MAC_addr	nil	C6	gateway	C4
C2	ip_addr	nil	C7	protocol/port	C1,C2
C3	login	nil	C8	proxy_http/socks	nil
C4	subnetting	C2	C9	proxy_connection	C3,C4,C7,C8
C5	route_selection	C4,C6	C10	connection	C2,C3,C5

In Fig.1 some of knowledge items can have a low-level model. With such hierarchical models, we might illustrate how to compute the knowledge items that has not mastered. If there is an evidence that a student fails to answer questions about C5,C9, using Reiter's idea about consistency diagnosis to compute conflict sets and hitting sets. We get conflict sets as {C2,C4,C5,C6} and {C1,C2,C3,C4,C7,C8,C9},

from the conflict set we can get hitting sets {C2}, {C4}, {C1,C5}, {C1,C6}, {C3,C5}, {C3,C6}, {C7,C5}, {C7,C6}, {C8,C5}, {C8,C6},{C9,C5},{C9,C6}. There are 10 results, but which one is the most possible to conduce the evidence?

3.2 Analysing Results of MBCD

We consider that apply probability statistics to sequence with average probability of each hitting set. In order to keep the accuracy of the diagnostic system, we select the total of hitting sets (which has N subsets) as $\lceil \log_2 N \rceil$. Table 2 displays the relationship between knowledge items of compute Network and questions (Q1, Q2, Q3,…, Q10). Each value (Qi, Cj) ranges from 0 to 5, represents the relationship between Qi and Cj. 0 indicates no relationship; ERROR(j) denotes the total failure of concept Cj for all students; pro(Cj) = ERROR(j)/(count of total students) represents the correlation degree of Cj to the incorrect answers. In this paper, making 30 students to do the test questions. For example, there are eight students has answer wrong at Q1, and the total number is 30, so the pro(C1)=8/30=0.27.

Table 2. Correlation degree of knowledge items with wrong answer

	Cj									
Qi	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
Q1	5	0	0	0	0	0	0	0	0	0
Q2	0	5	0	0	0	0	0	0	0	0
Q3	0	0	5	0	0	0	0	0	0	0
Q4	0	0	0	5	0	0	0	0	0	0
Q5	0	0	0	0	5	0	0	0	0	0
Q6	0	0	0	0	0	5	0	0	0	0
Q7	0	0	0	0	0	0	5	0	0	0
Q8	0	0	0	0	0	0	0	5	0	0
Q9	0	0	0	0	0	0	0	0	5	0
Q10	0	0	0	0	0	0	0	0	0	5
ERROR	3	9	0	12	16	15	10	4	24	21
Pro(Cj)	0.1	0.3	0	0.4	0.53	0.5	0.33	0.13	0.8	0.7

Before what we will do, we could do some deal with the hitting sets:(1)Check each wrong knowledge that wether all its required knowledge items are right, if it is so, then we consider that its fault is irrelevant with any other knowledge items. The conflict set on it is only the wrong knowledge item itself; (2)Check each wrong knowledge item Ci, wether the knowledge items whose required knowledge item is and is only Ci that is wrong, if not, we consider that Ci is wrong because of careless. And will not compute conflict set and hitting with it. For example, the probability of {C1,C5} is Pro(C1)+Pro(C5)/2, and the "Rank" is the sequenceing of the results. From Table 3, there ten subsets, the number will be selected is $\lceil \log_2 12 \rceil$=4, we can see that the best result is {C5,C9}, and others are {C9,C6}, {C5,C7},{C7,C6}.

Table 3. Suspicious degree of each hitting set

resluts	probability	Rank	*resluts*	probability	Rank
{C2}	0.3	10	**{C7,C6}**	**0.5+0.33/2=0.415**	**4**
{C4}	0.4	5	{C7,C5}	0.53+0.33/2=0.43	3
{C1,C5}	0.1+0.53/2=0.315	8	**{C8,C6}**	**0.5+0.13/2=0.315**	**7**
{C1,C6}	0.1+0.5/2=0.3	9	**{C8,C5}**	**0.53+0.13/2=0.33**	**6**
{C3,C5}	0+0.53/2=0.265	11	**{C9,C6}**	**0.5+0.8/2=0.615**	**2**
{C3,C6}	0+0.5/2=0.25	12	**{C9,C5}**	**0.53+0.8/2=0.665**	**1**

4 Experimental Results

In this section we present our experiments that show the performance of the MBCD in our system. We have implemented the MBCD in Smalltalk using VisualWorks 7.3 non-commercial, and it acts as a model-based diagnosis engine [9,10] that uses either a theorem prover or a constraint propagator to find conflict sets and thus to make diagnosis by using algorithms for computing hitting sets [6]. In our experiments, the MBCD runs on a PC (Pentium(R) Dual-Core E5200, 2.50GHz CPU speed, 2G memory, running on Windows XP), and works together with a testing system running on a server (450MHz CPU speed, 2G memory, running on Linux), which provides exam questions in Web pages and checks students' answers.

4.1 Diagnosing Students' Test Performance

In our experiment, we have 20 students who are new to Computer Network which has been manually modeled like Figure 1. To evaluate the MBCD's value, we put them into two groups, each with 10 students. The group who uses the MBCD is called group A, while the other is called group B who are not assisted with the MBCD and are expected to be capable of self-learning.

Before the testing, all students take 8 hours to study the knowledge items in the Computer Network course. Then they are asked to test themselves with questionnaire containing exam questions generated by the testing system according to what they have learned. Each test paper contains 10 exam questions, either single choice questions or multiple choice questions, each concerning about the knowledge items that students are expected to have acquired so far.

4.1.1 First Testing

Table 4 shows the test performance of students in group A, the "Score" is student's points obtained in testing, the "Wrong" is knowledge items on which the student fails, the "Conflicts" and the "Diagnosis" respectively refer to the conflict set and the hitting set, both computed by the MBCD, the"Analysis" is the results that using suspicious degree to deal with the hitting sets. Table 5 shows the test performance of students in group B.

Table 4. First test performance of group A

Who	Score	Wrong	Conflict	Diagnosis	Analysis
A1	100	Nil	Nil	Nil	Nil
A2	90	C2	{C2}	Nil	careless
A3	90	C3	{C3}	{C3}	{C3}
A4	70	C4,C5, C6	{C2,C4},{C2,C4,C6}, {C2,C4,C6,C5}	{C2} {C4}	{C4}
A5	80	C2,C4	{C2,C4}, {C2}	{C2}	{C2}
A6	70	C4,C6, C7	{C2,C4}, {C1,C2,C7}, {C4,C2,C6)}	{C2},{C4,C1}, {C4,C7}	{C2} {C4,C7}
A7	70	C7,C8, C9	{C1,C2,C7},{C8},{C1,C2,C3,C4,C7, C8,C9}	{C1,C8},{C2, C8},{C7,C8}	{C7,C8}, {C2,C8}
A8	90	C5	{C2,C4,C5,C6}	{C2},{C4}, {C5},{C6}	{C5}
A9	60	C7,C4, C9,C10	{C2,C4},{C1,C2,C7},{C1,C2,C3,C4, C7,C8,C9},{C2,C3,C4,C5,C6,C10}	{C1,C4},{C2}, {C4,C7}	{C4,C7} {C2}
A10	80	C8,C9	{C1,C2,C3,C4,C7,C8,C9},{C8}	{C8}	{C8}

In Table 4, the group's average score is 74, in particular, A1's 100 points suggests that A1 has almost mastered all knowledge items, so A1 is excluded from the second examination. For A2, because C4 depends entirely on C2 is right, so say that C2's failure is for careless, then A2 also do not go in for the second examination.

In Table 5, we take for granted that there is no other reason but the student has some difficulties in the failed knowledge items. The group's average score is 74, student B1 is excluded from the second examination. In the second examination, group A has 8 students present and group B has 9 students present.

Table 5. First test performance of group B

Who	Score	Wrong	Reason	Who	Score	Wrong	Reason
B1	100	Nil	Nil	B6	70	C4,C6,C7	{C4,C6,C7}
B2	90	C2	{C2}	B7	70	C7,C8,C9	{C7,C8,C9}
B3	90	C3	{C3}	B8	90	C5	{C5}
B4	70	C4,C5, C6	{C4,C5, C6}	B9	60	C7,C4,C9, C10	{C7,C4, C9,C10}
B5	80	C2,C4	{C2,C4}	B10	80	C8,C9	{C8,C9}

4.2 Second Time for Experiment

Before starting the second testing, all students are allowed to spend another 4 hours on learning. Then both groups take the second examination. The second result of students in group A shows that their average score is increased from 74 to 92, and students A3, A5, A8, A10 have masted all knowledge items. And group B's result is that their average score is 89 (vs. 74 in the first testing). Whereas student B2,B3,B8 achieves 100 points. Comparing with the group A which has 6 students completed the course, group B only has 4 students. It means that the students using our diagnostic system could increase learning efficiency.

This experience shown that the student use the diagnostic system proposed by us would see what the essential reasons for the knowledge he/she does not master, and give him/her some suggestion to help to determine the next step what to learn first.

5 Conclusion

This paper considers the problem of cognitive diagnosis of students' test performance as an application of model-based diagnosis, we present an approach on model-based cognitive diagnosis by running a logical representation of a course against the student's test performance in order to pinpoint their faults in cognition and to improve their learning performance. Experimental results show that the group of students with such understanding can improve performance greatly in an E-learning environment.

We are working on interactive cognitive diagnosis from multiple observations, and how to extend our course modeling with fault mode describing that the learning components tend to fail in specified ways. Motivated by course configuration in [3], we also consider how to enhance our MBCD system through test planning.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (No.60775028), the Major Projects of Technology Bureau of Dalian (No.2007A14GXD42), and the Dalian Maritime University Backbone Youth Foundation(NO.2011QN033, No.2009JC29).

References

1. Ohlsson, S.: Some principles of intelligent tutoring. Instructional Science 14(3), 293–326 (1986)
2. Huebner, A.: An Overview of Recent Developments in Cognitive Diagnostic Computer Adaptive Assessments. Practical Assessment, Research & Evaluation 15(3), 1–7 (2010)
3. Xu, X., Chang, H., Douglas, J.: A simulation study to compare CAT strategies for cognitive diagnosis. In: The Annual Meeting of the National Council of Measurement in Education, Chicago (April 2003)
4. Dillenbourg, P.: A framework for learner modeling. Interactive Learning Environments 2, 111–137 (1992)
5. Self, J.: Model-based cognitive diagnosis. User Modeling and User-Adapted Interaction (1993)
6. Reiter, R.: A theory of diagnosis from first principles. Artificial Intelligence 32(1), 57–95 (1987)
7. Fleischanderl, G., Havelka, T., Schreiner, H., Stumptner, M., Wotawa, F.: DiKe - A Model-Based Diagnosis Kernel and Its Application. In: Baader, F., Brewka, G., Eiter, T. (eds.) KI 2001. LNCS (LNAI), vol. 2174, pp. 440–454. Springer, Heidelberg (2001)
8. Greiner, R.: A correction to the algorithm in reiter's theory of diagnosis. Artificial Intelligence 41(1), 79–88 (1989)
9. Felfernig, A., Felfernig, E.A., Friedrich, G., Stumptner, M., Jannach, D.: Consistency-based diagnosis of configuration knowledge bases. In: Development. Harvard University Press Goddard, Prentice Hall, Cambridge, Mass (2004)
10. Jiang, Y.-f., Lin, L.: Computing the minimal hitting sets with binary hs-tree. Journal of Software 13(22), 2267–2274 (2002)

Design and Implementation of Locust Data Collecting System Based on Android*

YueQi Han[1], Xunan Shan[2], Dehai Zhu[3], XiaoDong Zhang[3], Nan Zhang[1], Yanan Zhang[1], and Lin Li[1,**]

[1] China Agricultural University, College of Information and Electrical Engineering, Computer Science and Technology, TsingHua DongLu Street No.17, Beijing, 100083
lilincau@gmail.com
[2] National Agro-Technical Extension and Service Centre, No.20 building, Maizidian Street, Chaoyang district, Beijing, 100125
[3] China Agricultural University, College of Information and Electrical Engineering, Geography Information System, TsingHua DongLu Street No.17, Beijing, 100083

Abstract. This Paper is mainly focused on the Locust Data Collecting System based on Android platform, and designs the structure of data acquisition and data transmit system. The system includes data acquisition unit, data transmission unit and process unit in the server. In addition, the paper discusses the implementation of the Data Collection in mobile and the realization of the programming. The experiment indicates that the method is feasible and effective.

Keywords: Locust Data Collecting; Data transmit; Android; Collection in mobile.

1 Introduction

The plague of locusts is one of the fatal disasters in Agriculture, which composites three main natural calamities with drought and flood. The plague of locusts was considered as a frequent consequence in the history of China, mainly caused by oriental migratory locust[1]. The oriental migratory locusts are prolific breeders, with short development time, meanwhile they are omnivorous, wide spread, and congregate, which makes a great contribution to the plague of locusts. Therefore, effective method on locust prevention and detection is very important in this issue. Locust data collection in the field and real-time data transmission is the basis of locust prevention, detection and research. In this paper we provide design and realization of Locust data collection system on smart phone base on the information of the field work.

In accordance with the business requirement of locust data collection, the system should have features as follow. First, we need to choose the dominant smart phone

* Special Fund for Agro-scientific Research in the Public Interest, and "Green Technology Research and Demonstration on Prevention and Control of Migratory Locusts in P.R. China", which is sponsored by Public Service Sectors (Agriculture) Special Research under grant 200903021.

** Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 328–337, 2011.

whose operating system has a big market and development potential. Second, the user can be divided into county level, province level, prefecture city level and ministerial level. The county level users mainly accomplish the collection of locust data and report of the data, while the others have their corresponding rights for searching the infection of the plague of locusts in their own unit. The county level users collect the locust information base on the area group by the plague of locusts. So the system requires the information of locust area from the data center and then exquisite the information. Third, since the signal is unstable, we should save the data when the signal is weak and transmit the data when the signal turns well. Finally, the collected data will include the sample's GPS information and photos. How to make the data acquisition more continent and finish the great data communication will be a hang-up. This paper introduces a practical solution to these problems.

2 Structure and Function of the System

The system will be developed on Android platform according to the self-requirement. The target is to complete the system on basis on of the hardware in smart phone, which helps the data collecting in the field more effective, data acquisition and data real-time transmission more convenient. Android mobile operating system is based on a modified version of the Linux kernel. Google and other members of the Open Handset Alliance collaborated on Android's development and release. The android operating system is the world's best-selling smart phone platform and wins the support from a lot of cellular telephone manufacturers. The platform has a good support for the communication networks and the next generation network (3G), realizes the data transmission, and returns with web services. Extricating the collector from the complex equipment, the system integrates the GPS function and camera function which are provided by the API interface in the Android platform. In particular, different functions with this system are specified for the proper authorized users, so that the precision of the sample is greatly improved. Finally, the system provides the function of inquiry for the collectors inquiring the historical data of the samples and the function of query the characteristics of locust which provide convenience to the collector in the field. This paper introduces functions and mechanism of the system in detail.

2.1 Function of Locust Data Collecting System

Locust Data Collecting System is a sub-system of the locust digital information platform. The function of the whole system is shown in Figure 1. The system functions consist of user management, information management in locust area, information management in obstacle, information management in infection of locust, information management in prevention, query of the characteristics of locust. The user management consists of login verification and data transmission verification, which provides an integrated safety service for the system. Information management in locust area inputs the edge information, collects the information of the locust area on one side, and provides the collectors with locust area information, so that it will be easier for the collectors to get and query the information. Information management in obstacle collects the information of obstacle, which provides references for the follow-through

such as aircraft navigation and mechanical spray. Information management in infection is key function of the system, which is applied for collect the locust data information by the input of UI. Information management in prevention is applied to input the method in the prevention and get the effect of the prevention. Query of the characteristics provides the information about the characteristics of locust, so that the collectors can inquiry the information about the locust's type and habitat in a convenient way.

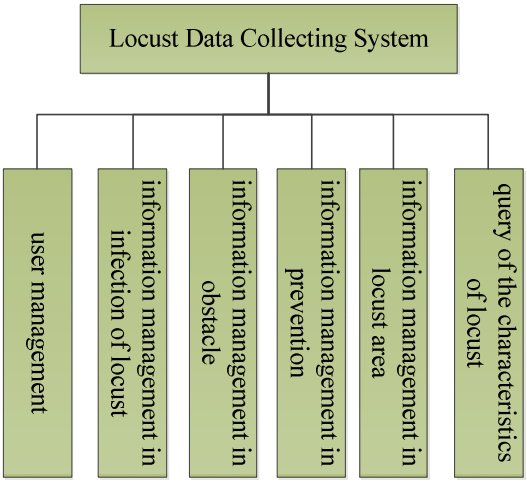


Fig. 1. The function of the whole system

2.2 Mechanism of the System

Mechanism of the system design is shown in Figure 2. First, capture the precise position information by using inner A-GPS in smart phone, take photos with inner camera of the phone, then form integrated information of the sample by the UI. Second, take the integrated information to the data transmission module and realize the encapsulation of the information, then transmit the encapsulated information to the server by web services. Finally, application in the server will process the information; post the processed information to data center and return the result to the client; the steps realize the real-time data transmission. The paper will give a detailed introduction of the three units.

First and foremost we discuss the data acquisition unit. This unit realizes the function of information collection and data transmission. Design of this unit is shown in Figure 2. Data acquisition unit collects different types of information, such as GPS information, infection of locust information, prevention information, and obstacle information. Furthermore, this unit provides the inquiry of the collected data and characteristics of locust. The system completes the capture the photos by inner camera. Sometimes the data cannot be uploaded to the server immediately because of the sick signal. As a result, we should add a local data cache to the system to ensure the information storage and data transmission when the network is well. The local data cache can be realized by the mobile database.

The following one to be introduced is data transmission unit. The process of the data transmission is that it transmits the data (provided by data acquisition) unit to the server in time and the server return the data back to the client. The data transmission unit is shown in the Figure2. This unit has two parts: the conversion of data and the communication between the server and the client. It makes the transmission more efficient that the system encapsulates and transforms the data by the format of the transport. And the transmission of the pictures also needs a data transformation. The interaction between client and server is that the client remote calls the web services. The network between client and server in this system includes GPRS and WCDMA. GPRS, short for General Packet Radio Service, is a wireless technology that allows the smart phone user to quickly connect to the network and obtain good data rates. GPRS Tunneling Protocol (GTP) is a group of IP-based communication protocols used to carry General Packet Radio Service (GPRS) within GSM and UMTS networks. The theoretical transmission rate is 171.2kbps, yet the actual transmission rate is about 40kbps. WCDMA (Wideband Code Division Multiple Access) is the radio access scheme used for 3rd generation cellular systems, which supports wideband services like high-speed Internet access, video and image transmission with quality as high as the fixed networks.

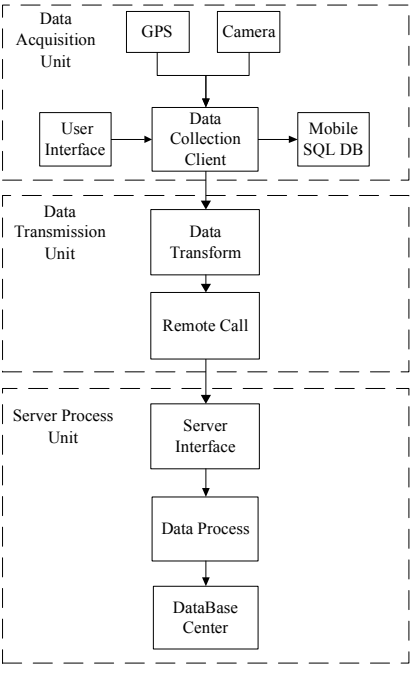


Fig. 2. Mechanism of the system

WCDMA standard was created by 3GPP (Third Generation Partnership Project) to ensure interoperability among different 3G networks. The static transmission rate is 2Mbps. Therefore, the system could call the remote methods to transmit the data between client and server when the smart phone accesses the GPRS or WCDMA.

Finally, I will give the description to the last part——process unit in the server. The unit in server is to process the data uploaded by the client and return a value to the client. If the data is correct, it will be preserved in the database center. The key point of the process is to analyze the uploaded data and check whether the data fays in with a certain rule or not. If it follows the rule, the data will be preserved in the database center and the client will get a successful message; or, an error message.

3 Design and Implementation of the System

Based on the information above, we choose the emulator and HuaweiU8500 as the client for the system. The main interface of the system is shown in Figure 3.The perspective of technology should be taken into account during design and implementation of the system. This paper discusses the key techniques and difficulties of the system, and gives the solution of those difficulties. These solutions make the feature of the system, and a detailed account of the solutions will be given as follow.

Fig. 3. Main interface of the system

3.1 The Design and Implementation of the Database

Android provides APIs for the developers, and some of the API involves the SQLite, so that the developers can easily add, delete, modify data, create and drop the table or database by the APIs.

We use the system API to access the database and create an entity/object to manipulate the database, the class named SqliteOpenHelper. The object concludes the manipulate method to the database, for instance, create a database, dropdown a database, modify a database or change the version of a database. Then we define the

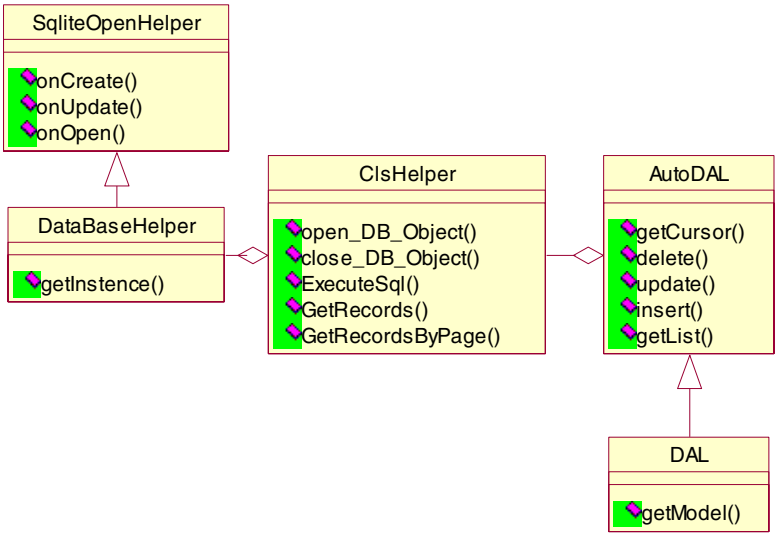


Fig. 4. Class diagram of SQLite access

object as abstract class, and implement the abstract methods in the super class——DataBaseHelper. In order to modify the method more easily, we create a class named ClsHelper which aggregates the DataBaseHelper. Then we program the table manipulate method in it, for example, add, delete, modify data, and the method are general method which cannot be called by system directly. Next, we create the AutoDAL class which aggregates the ClsHelper to create for each table in the database. AutoDAL provides the methods which can be directly called by system. And then, DAL which inherits class AutoDAL is created. The DAL generates the methods for the tables with different parameters. The full class diagram is shown in Figure 4.

3.2 Data Synchronization Technology in the System

In order to ensure the requirement of real-time characteristic, the system should perform the real-time data transmission. Data synchronization technology can be parceled into two parts: the server and the client.

There are two parts in the process on the client, one is data transformation, and the other is the transmission between server and client. Data transformation is a process that encapsulates the information to be delivered. XML takes great advantage in data transmission on the Web. It is an extensive language and can be read and written in data flow. Web Services are modified in the data transmission. The Web Service method will be called in the client and the research on the software development is based on SOA. So we should provide a chart of the process above. The class diagram is show in Figure 4. First, we define the interface ITable and declare two abstract methods in it; the methods are used for data transmission, and Implement the method in XML_Manager class; XML_Manager declare the general method to the basic data transformation. And then, we create corresponding classes which inherit XML_Manager (we name them Table_XML_Manager in Figure 5) for the

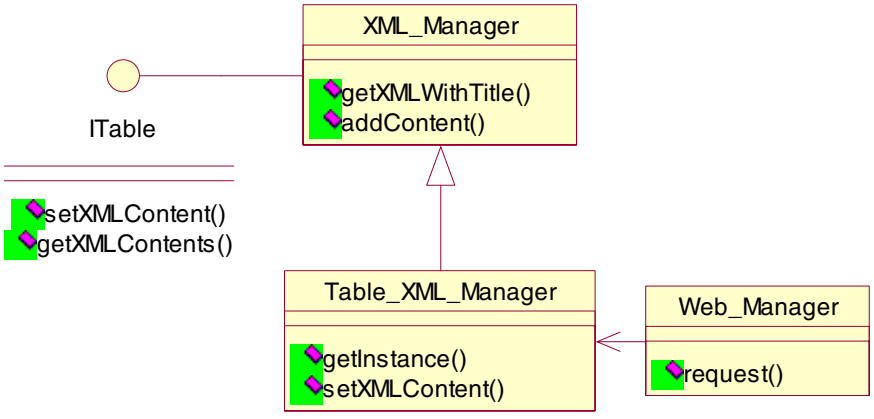


Fig. 5. Class diagram of data transformation and transmission

transmission of each table in database. Finally, we use the Web_Manager to invoke the Web Service and transmit the packed XML data as a parameter of Web Service method. Furthermore, the pictures should be transformed into bit streams in the transmission and the service transforms the bit flow into picture file, and then keeps the path as the attribute values into the database.

For the benefit of the efficient of data synchronous, we create a copy of each table when the system transmits the data. One of them is used to keep the historical data while the other is used to hold the data to synchronous. The table which holds the synchronous data will delete the data when synchronization is succeeded, so that the data will not duplicate synchronous, meanwhile it will take corresponding measures to solve the problems when the system gets the error information. This design can effectively reduce data uploading time.

The functions of the server include Web Service method, data inspecting and data submitting to the center database. Based on Visual Studio 2008 development, the unit in server uses the Web Method scope in Web Service, so that the client can access the server by SOA (shown in Figure 6). We should define different method to deal with different XML data flow.

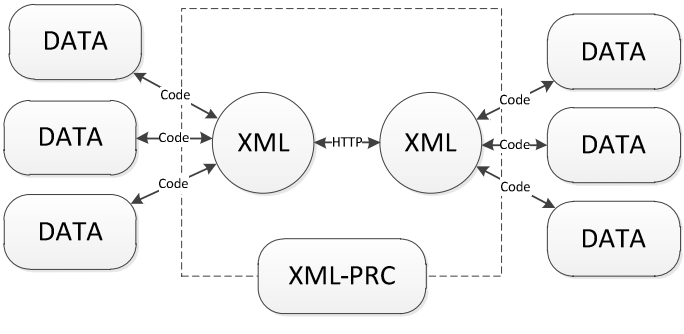


Fig. 6. Call of Web Services

Meanwhile the server needs to analyze the flow. Then, the server determines whether the analyzed data fays in with the rule or not. If it follow the rule, it will be preserved in the database center and return a successful message; or, return an error message.

Besides, the system takes the singleton mode to access the server database, which makes the access of the database safer and the speed higher.

3.3 GPS Data and Image Collection

Data acquisition core is the center of the system in the client, and the unit is developed on the Android platform. Based on the Eclipse, the unit uses Java programming language. In accordance with business process, the unit is separated into several activities to interact with users. The GPS data is captured by the process which runs in background, and Android platform provides the way of the process named service. The GPS service is started by calling the startService() method when the user successfully login, and the GPS data is sent to the activity by the GPS Service when it is captured. After that the data will be shown in the User Interface. The API interface calls the function of camera, and returns the image information by data flow. The function in the process is directly shown as Figure 7.

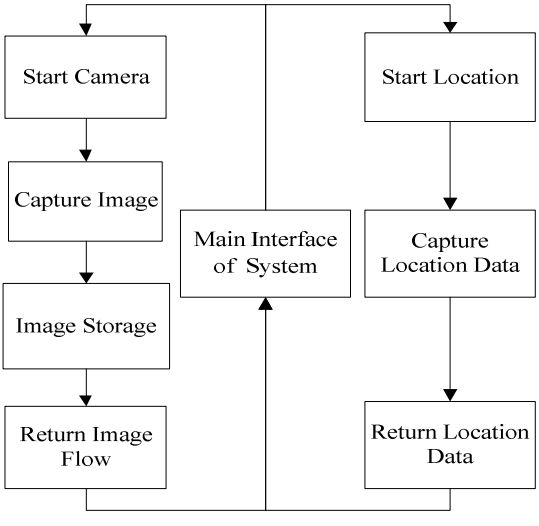


Fig. 7. The function diagram of collection in GPS data and image

3.4 User Authentication and Information Downloading

User authentication is composed of login verification and transmission verification. Login verification includes the local check and the internet check. Local Verification inspects user's information by the record in the local database. By this way, users are

permitted to work normally when the signal is weak, while the internet verification inspects the user's information on the server by SOA. Because of the website in the program, the information of users may be changed by the web. So the user's login should be checked by internet for the new information, for example, the password. The information will be updated automatically when the user information is checked by the internet. But when the data transmission is vivificated, the user information must be checked by internet. Furthermore, the information about the locust area can be downloaded. The information helps to locate areas accurately and ensure the relationship between the locust area and the sample. The whole process of verification is shown in Figure 8.

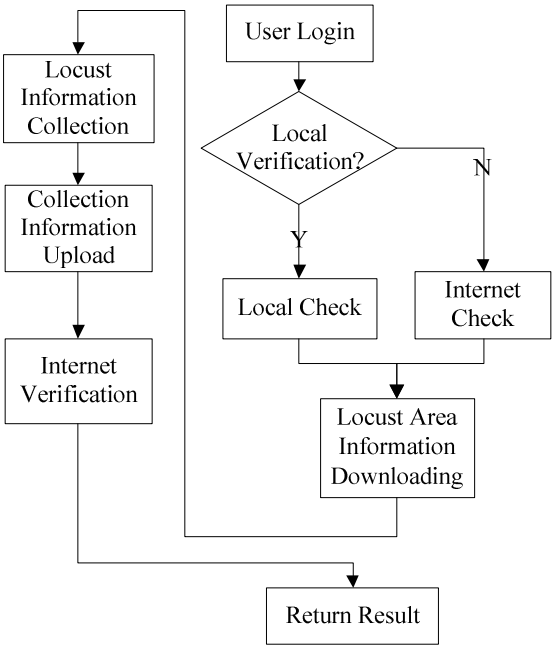


Fig. 8. The flow chart of login verification

4 Conclusion

This paper is concentrated on the research of the Locust Data Collecting System based on the Android platform, and designs the structure and mechanism of the system. A detailed introduction to the frame and the structure including data acquisition unit, data transmit unit and process unit in the server was shown. The implementation and method of the system, realization of the system by programming were also discussed. The technology in Locust Data Collecting System is a general application in the data collection field, yet it is a significant reference for the solution of similar issues.

References

1. Wan, X., Zhu, D., Zhang, L.: Research and development of field data collecting system of locust. Agricultural Network Information (1), 22–25 (2006) (in Chinese with English abstract)
2. Mu, L., Liu, G., Huang, J.: Design of farm field data collect ion and transmission system based on Java phone. Transactions of the CSAE 22(11), 165–169 (2006) (in Chinese with English abstract)
3. Li, L., Li, M.: Java Mobile Phone Oriented GPS Data Picking and Wireless Transmitting System. Geomatics and Information Science of Wuhan University 29(8), 728–731 (2004) (in Chinese with English abstract)
4. Android Developers, `http://developer.android.comlindex.html`
5. EoeAndroid, `http://www.eoeandroid.com/`
6. Whipple, J., Arensman, W., Boler, M.S.: John Whipple William Arensman Marian Starr Boler. In: Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA (October 2009)
7. Lai, Y.C., Han, F., Yeh, Y.-H.: A GPS Navigation System with QR Code Decoding and Friend Positioning in Smart Phones. In: 2nd International Conference on Education Technology and Computer, ICETC (2010)
8. Wu, Y., Luo, J., Luo, L.: Porting mobile web application engine to the Android platform. In: 10th IEEE International Conference on Computer and Information Technology, CIT 2010 (2010)
9. Noguera, J.M., Segura, R.J., Oga'yar, C.J., Joan-Arinyo, R.: Navigating large terrains using commodity mobile devices. Computers &Geosciences (November 9, 2010) (article in press)
10. Hui, S.C., Fong, A.C.M., Jha, G.: A web-based intelligent fault diagnosis system for customer service support. Engineering Applications of Artificial Intelligence (14), 537–548 (2001)

Verification of Cutting Zone Machinability during the Turning of a New Austenitic Stainless Steel

Jozef Jurko[1], Anton Panda[1], Mário Gajdoš[2], and Tadeusz Zaborowski[3]

[1] Technical university in Košice, Faculty of Manufacturing Technologies,
Štúrova 31, 080 01 Prešov, Slovakia
[2] ELBA a.s. Kremnica, ČSA 264/58 967 01 Kremnica, Slovakia
[3] Institute for Scientific Research and Expertises,
66-400 Gorzów Wielkopolski, Łokietka 29, Poland
{jozef.jurko,anton.panda}@tuke.sk,
mgajdos@elba.sk, tazab@sukurs2.pl

Abstract. In this study presents the conclusions of machinability tests on a new austenitic stainless steel X5Cr18Ni9MoTiN, which applicated in food processing industry, and describes important concurrent parameters for the cutting zone during the process of turning. The machinability of stainless steels is examined based on the cutting tests, cutting speed presented by v_c-T diagrams. The effect of cutting speed are analysed by chip formation and tool wear mechanisms. Based on the cutting tests, cutting speeds of 40 to 200 m/min, feed rate of 0.04 to 1.2 mm and solid carbide indexable insert CNMG 120404-MF. Tool wear criterion, VB_K value of 0.3 mm. Chip were analysed by Semi Electron Microscopy (SEM).

Keywords: machinability, turning, cutting zone, stainless steels.

1 Introduction

This Precise and reliable information on the machinability of a material before it enters the machining process is a necessity, and hypotheses must be tested through verification of actual methods. This article presents conclusions of machinability tests on a new austenitic stainless steels X5Cr18Ni9MoTiN and describes appropriate parameters for the cutting zone during the process of turning. Austenitic steels are the most extensive and thus the most important category of stainless steels. To properly describe the cutting zone it is necessary to describe the regions and test parameters [3]: Primary plastic deformation zone, Secondary plastic deformation zone, Tertiary plastic deformation zone, Cutting surface, The gradually-deformed region of the cut layer.

2 Evaluation and Experimental Results

This article concerns itself with the evaluation of selected domains of machinability in compliance with EN ISO 3685 standards [11]. The experiments were performed in

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 338–345, 2011.

laboratory conditions and verified in real conditions during manufacture. The set-up contained the following components: a VMF-100 CNC machining centrum a CNMG 120404-MF cutting tool with M20 cutter, and PCLNR 2525M12 tool holder.

Table 1. Chemical composition of austenitic stainless steels

Chemical element	Chemical Composition in *wt* [%]
C	0.05
Cr	18.0
Ni	9.0
Mo	1.0
Ti	0.06
Si	0.6
N	0.07
P	0.04
S	0.04

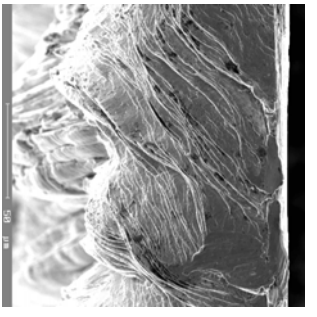
The materials to be machined were type of a new austenitic stainless steels with chemical composition listed in Table 1. The dimension of each piece was of diameter (*d)* of 120 mm and length (ℓ) of 200 mm. The cutting process employed was axial dry machining (DM), and the cutting speed was defined at intervals of v_c=40 to 200 m/min, the feed was advanced from intervals of f=0.04 to 1.2 mm per rev., and cutting depth a_P=0.5 mm.

Researching the cutting zone (the interaction between the tool, the workpiece, and the chips) is to capture its state at the moment of the creation of the chip (the so-called root of the chip), shown in Figure 1. Figure 1 is a model of the cutting zone, which illustrates the underlying phenomena in the separation of the cutting from the material of the workpiece. It is important to define the shear level in the cutting zone. More autors states, that the depth of the shear level follows the formula $0.05h \leq h_{SP} \leq 0.1h$, where h is the thickness of the cut section and h_{SP} is the depth of the shear level. The size of this local region was determined through the help of electron microscope analysis, and the results are displayed in Table 2.

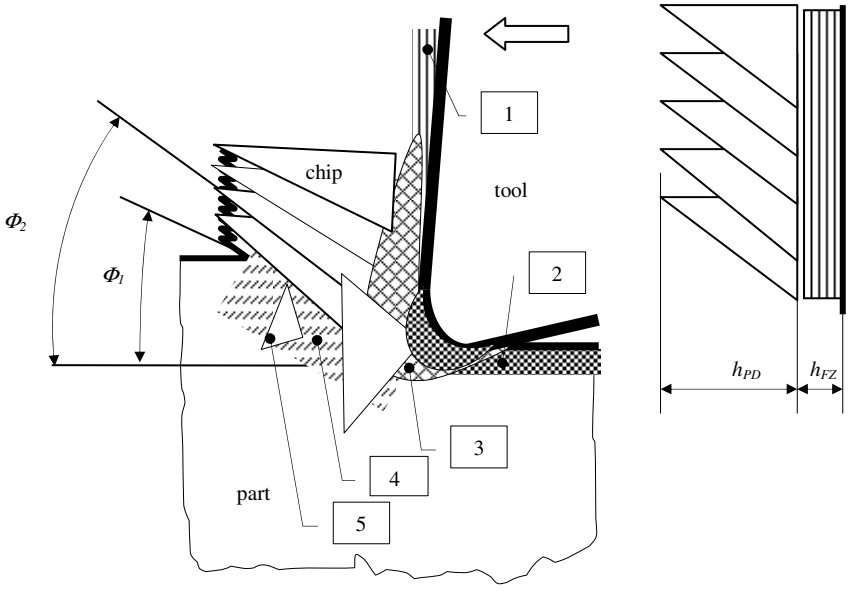
Another phenomenon in the cutting zone is the marked destruction of the X5Cr18Ni9MoTiN steel in front of the cutting edge, as shown in Figure 2a. The Figure 1 shows the project of the model of chip production in the cutting zone and it is the result of the analysis of the examined cutting zone. The result of the cutting zone evaluation for cutting conditions (v_c=150 m/min, a_p=0.6 mm a f=0.1 mm per rev.) the defining of the shear plane angle Φ_1 and the texture angle Φ_2. For the X5Cr18Ni9MoTiN steel, the angle is Φ_1=25 to 26° the angle Φ_2=34 to 35°. The important values of size of these areas are: the area of the plastically deformed material (chip), the flow zone area and the area of the hardening of the machined surface.

According to Cook [1], [10] material machinability is a quality of the material that expresses its capacity to process the work piece from the point of view of its functional qualities. According to [7], [9] material machinability is expressed as a quality of the material, which is defined by the state of the cut surface, the creation and shaping of chips, the effect of cutting forces and the tool life of the cutting edge. According to [4], material machinability is a quality of the object material, which

expresses its qualitative state by yielding to the effect of the cutting wedge. Austenitic Cr-Ni steels are, as a result of their higher ductility, more prone to surface strain hardening, which compared to construction steel can be up to 1.5 times as great by [4]. In a non-deformed state austenitic steels are not as hard as C45 steel, but in cases of great deformation they are greatly harder than ferritic-perlitic steel by [4], [6]. Sintered carbide tools are not as sensitive to temperature on touching plates as high-speed steel, and can be used to attain higher performance, but in this case they have greater pressure stress, which directly influences the process of adhesive wear by [8].



a)



b)

Fig. 1. Cutting zone and model of chip formation during turning of steel X5Cr18Ni9MoTiN a-chip texture, magnification 1900x, b-chip production model, 1-Flow Zone 2-stress hardening of the surface at the point of cutting, 3-volume of driven material, 4-slippage region, 5-beginning of creation of solid segments of material, (Φ_1-angle of the shear level, Φ_2-chip angle)

The criterion VB_K=0.3 mm was applied during evaluation. The cutting process conditions were designed based on the needs of the material and on the operation of the finished surface. The results of the long-term test after exhaustive analysis of selected data are show on Figure 3. The results of this method is Taylor equation $T = \frac{29{,}63.10^3}{v_c^{1,4149}}$, where the values shown are appropriately acquired values from analysis of measured values, Figure 3.

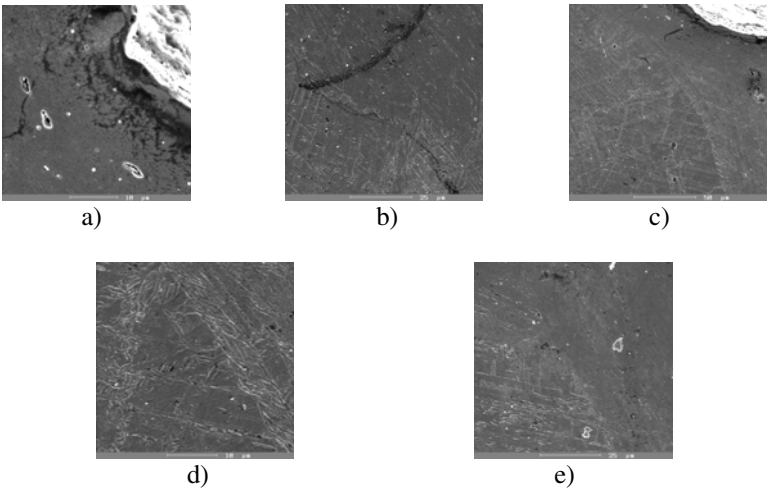


Fig. 2. Accompanying phenomena in the cutting zone for the X5Cr18Ni9MoTiN steel, a-texture of the chip, magnification 1900x, b-primary zone of plastic deformation, magnification 1900x, c- plastic deformation – slip field, magnification 750x, d-slip lines, magnification 3800x, e-limit between the deformed and non-deformed material, magnification 1900x

Table 2. Cutting Zone Parameters

Parameters	Result
h_{PD} [mm]	(75-88 %).h_t
h_{FZ} [mm]	(19-16 %).h_t
h_h [mm]	0.15
h_{SP} [mm]	0.9-1.0

where h_t is the width of cutting chip, h_{FZ} is the depth of the Flow zone, h_{PD} is the depth of the plastically-deformed material, h_h is the depth of the hardened machined surface and h_{SP} is the depth of the shear layer.

According to [2], [5], [10] in axial turning final force F breaks down to $F=F_c+F_p+F_f$. Constituent analysis was obtained with the aid of a tricomponent piezoelectric dynamometer KISTLER 9257A.

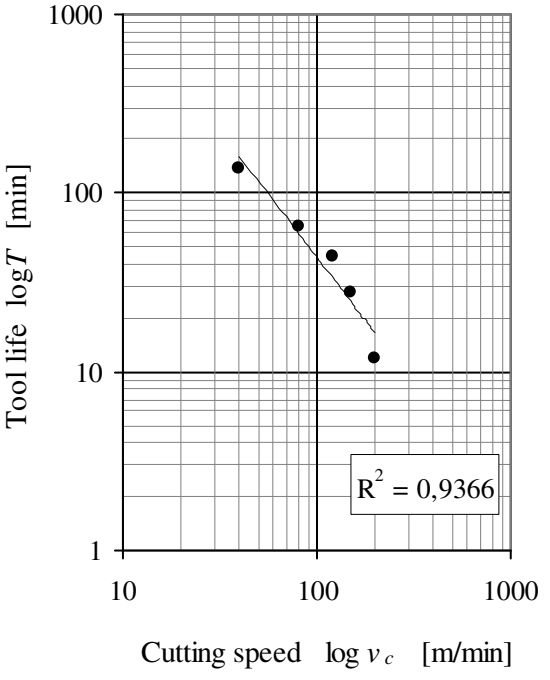


Fig. 3. Relation between the tool life of log T and cutting speed log v_c

The results are interesting from two standpoints: at smaller values of chip cross-section (feed values up to 0.4 mm) the specific cutting force was much higher, achieving values up to 22000 MPa. For greater values of chip cross-section (for feed values above 0.4 mm) the value of the specific cutting force peaks at 3500 to 7500 MPa. The BUE is one of the important factors which influences a change in the specific cutting force, which is formed with steel X5Cr18Ni9MoTiN mainly at cutting speeds between 55 and 65 m/min. For C45 steels this region is defined at cutting speeds of around 25 to 30 m/min. The final relation during turning according to this equation is stated in Table 3.

Table 3. Components of cutting force for stainless steel

Componential Equation for cutting force [N]		
$F_f = 15{,}81.10^6.a_p^{1{,}4632}.f^{0{,}2845}$	$F_p = 3411{,}35.a_p^{1{,}0822}.f^{0{,}4092}$	$F_c = 1988{,}25.a_p^{0{,}9736}.f^{0{,}4125}$

Evaluating the outer surface after turning and defining the cutting process conditions, the following parameters were used in the investigation: the outer surface roughness parameter *Ra* [μm] was measured on two measuring tools a HOMEL TESTER 1000 C, the micro hardness of the outer surface layer was evaluated following Vickers [HV]. The measured results are documented in Figure 4 and Figure 5.

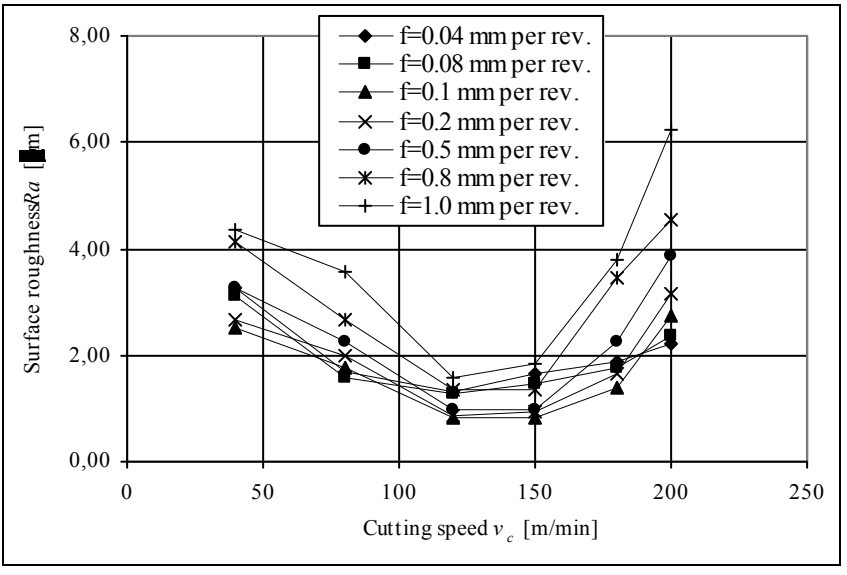


Fig. 4. Dependence surface roughness *Ra* on the cutting speed, X5Cr18Ni9MoTiN steel

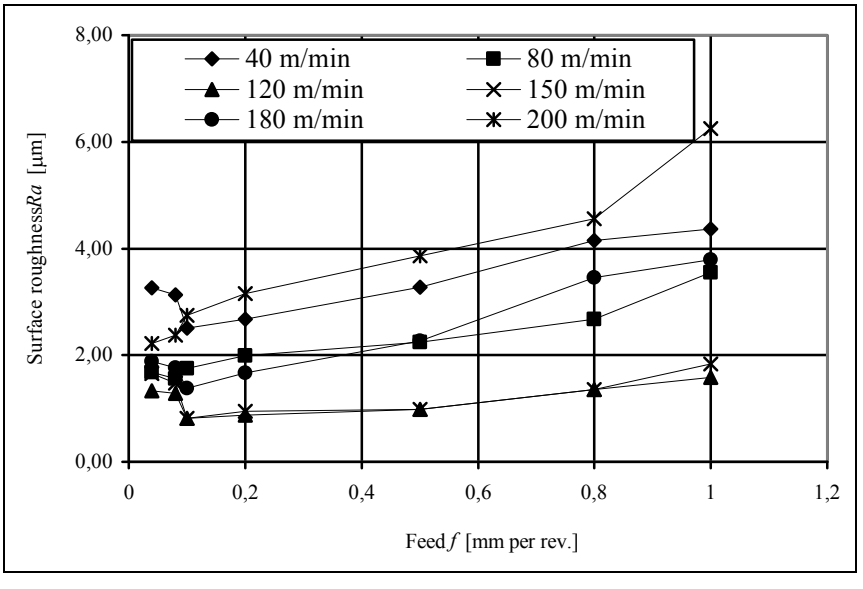
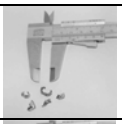
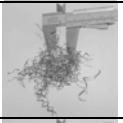
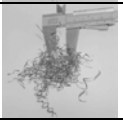


Fig. 5. Dependence surface roughness *Ra* on the feed, X5Cr18Ni9MoTiN steel

The acquired results are interesting in that for the defined conditions we can achieve a quality outer surface after cutting with roughness parameters down to around 0.8 µm. Very good results were mainly achieved when cutting speed was up

120 m/min to 150 m/min and the feed was interval up 0.1 mm per rev to 0.5 mm per rev.. The roughness value for the outer surface, *Ra*, reached around 0.8 µm. The process of forming chips by turning austenitic stainless steel is very problematic as the material is very tough, and the results of which are documented in intensive plastic deformation in the chips in Table 4.

Table 4. Forming of shavings in turning the X5Cr18Ni9MoTiN steel

Cutting speed v_c [m/min]	Depth of cut a_p [mm]	Feed rate f [mm per rev.]				
		to 0.04	0.04-0.08	0.08-0.1	0.1-0.5	0.5-1.0
40	0.8					
40-80						
80-120						
120-150						
150-200						

5 Conclusion

It is important for both theory and practical applications that essential conclusions come from measurement and analysis. The conclusions are as follows: defined tool life equation following Taylor, defined the equation for the cutting strength components, designed a model to generate chips, thermal analysis for the cutting process in the cutting zone, confirmation of surface strain hardening (change in mechanical properties) after cutting. Defined coefficients for kinetic machining of austenitic stainless steels, whereby X5Cr18Ni9MoTiN steel K_v=0.49-0.51, for C45 steel, the coefficient of kinetic machining was K_v=1.0.The machinability of austenitic stainless steel is two to three times worse that than for C45 on the basis of its chemical components: mainly chromium, nickel and other component elements. For the cutting process, it is necessary to use a tool that has a large cross section so that it can sufficiently dissipate the heat formed from the cutting zone. The wear of cutting tools may also affect the selection of appropriate geometry, mainly the positive angle

of the front of the tool as well as the required high surface quality of the tool's effective area. Avoid cutting interval speeds of 55 to 65 m/min, large feeds, lowered hardness in the cutting zones and low cutting speeds.

On the basis of experience, the authors recommend, for machining these types of steels, selecting criteria for automated production process based on the following order: For finished work, the criteria are set in the following manner: outer surface quality after cutting, forming of chips, kinematic processes and dynamic processes. This article is the result or much research work on the part of the authors in this field and the article presents actual conclusions that are currently being successfully implemented in machine shops.

Acknowledgments. The authors would like to thank in words the VEGA and KEGA grant agency at the Ministry of Education SR for supporting research work and co-financing the projects: Grant work KEGA #01/3173/2006 and Grant work VEGA #01/0048/2010.

References

1. Cook, N.: What is machinability. In: International Symposium on Influence of Metallurgy on Machinability, Deaborn Ohio, pp. 1–10 (1975)
2. Dautzenberg, J.H.: The Minimum Energy Principle for the Cutting Process in Theory and Experiments. Annals of the CIRP 32(1), 223–227 (1983)
3. Jurko, J.: Verification of cutting zone machinability during drilling of austenitic stainless steels. In: Proceedings of SPIE-The International Society for Optical Engineering 2008, Nanjing, vol. 7375(2) (2009)
4. Jurko, J., Gajdoš, M., Panda, A.: Accompanying phenomena in the cutting zone machinability during turning of stainless steels. International Journal of Machining and Machinability of Materials 5(4), 383–400 (2009)
5. Markulik, Š., Tompoš, A.: Overovanie technického stavu dvojkotúčovej brúsky Extol. Spravodaj ATD SR 1, 24–25 (2008)
6. Martinkovič, M.: Comparison of plastic deformation and grain boudaries orientation of steel. In: Dev. of Mat. Sci. in Res. and Educ., DMS-RE, pp. 50–51 (2009)
7. Mrkvica, I., Ochodek, V., Konderla, R.: Influence of a Tool Geometry on Surface Tension of a Workpiece. International Journal Manufacturing Technology 8, 76–79 (2009)
8. Suh, N.P.: New Theories of Wear and Their Implications for Tool Materials. Wear 62, 1–20 (1980)
9. Usui, E., Hirota, A.: Analytical Prediction of Three Dimensional Cutting Process. Part 2-Chip formation and Cutting Forces with Conventional Single-Point Tool. Journal of Engineering for Industry 100(2), 236–243 (1978)
10. Wei Yen, D., Wright, P.K.: A Remote Temperature Sensing Technique for Estimating the Cutting Interface Temperature. Journal of Engineering for Industry 113, 251–264 (1986)
11. ISO 3685: Tool life testing with single point turning tools

Improving Undergraduates' Experience of Online Learning: An Approach of Web-Mediated Self-Regulated Learning and Collaborative Learning

Chia-Wen Tsai

No.5 De-Ming Rd., Guishan Township, Taoyuan County 333, Taiwan, R.O.C.
Department of Information Management, Ming Chuan University, Taiwan
jawen12b@gmail.com

Abstract. Almost all undergraduates in Taiwan's universities and vocational schools have to take compulsory computing courses before they graduate. However, the teaching in computing courses commonly adopts traditional lectures with inappropriate examples. The author thus redesigned the web-mediated teaching methods based his reflection. This study involved 64 undergraduates received the interventions of web-mediated Self-regulated Learning (SRL) and Collaborative Learning (CL). The results of this study show that students had a better experience of course learning and more positive attitudes toward the interventions of web-mediated SRL and CL. Implications for teachers and schools are also provided in this study.

Keywords: Web-mediated SRL, Web-mediated CL, Application software education, Learning experience.

1 Introduction

Almost all undergraduates in Taiwan's academic universities and vocational schools have to take four to six compulsory computing courses before they graduate. However, the teaching in computing courses commonly adopts traditional lectures with inappropriate examples. The examples in different sections or chapters are not jointed or related. Students have very few opportunities to practice or independently solve problems with what they learned. Students who learn in this context may lose their learning motivation and involvement in this course, and even lose their future competence in the workplace [1]. Thus, collaborative learning (CL), which assists teaching a specific educational objective through a coordinated and shared activity by means of social interactions among group members is considered as an effective teaching method and was adopted in this study [2].

E-learning may pose a challenge for students to learn independently, without the teachers' on-the-spot assistance. Success in online courses often depends on students' abilities to successfully direct their own learning efforts [3]. King, Harner and Brown also hypothesize that self-regulation of learning is more important in the distance

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 346–349, 2011.

education context than the traditional context [4]. Therefore, the author in this study adopted self-regulated learning (SRL) to help students develop regular learning habits and improve their learning.

2 Empirical Study

2.1 Subjects

The subjects in this research were 64 undergraduates from a compulsory course titled 'Applied Information Technology: Networking'. Students came from the department of Security Management. The mean age of students was around 20 years old.

2.2 Course Involved

The course involved is a semester-long, 2 credit-hour course targeting second-year university students. The course content in the study consists of three parts: (a) Basic skills of using Internet and connections; (b) Conceptual knowledge of the guidelines for, and the possible mistakes and bugs in a business website; (c) Advanced computing skills of server installation, website architecture and planning, and web page programming.

2.3 Experimental Design

Students had to solve the same tasks but under different learning conditions. Students were divided into teams, each consisting of 6 to 7 members. Each team had to build up a quality website that met the guidelines in the literature and textbook. They also received instruction and had extra requirements for SRL.

2.4 Measurement

In this study, students were required to complete a short questionnaire to extract students' feedback. The author adopted the questionnaire of Boyle, Bradley, Chalk, Jones, and Pickard [5] to measure students' thoughts regarding blended course and interventions concerning web-mediated SRL and CL. The simple questions served as an instrument to gather data about students' thoughts regarding SRL, CL, and blended learning (BL). The results of this questionnaire are reported and presented with figures in this paper.

3 Results

In the questionnaire, students were asked what they thought regarding the intervention of CL. The results are summarized in Figure 1, and show that all of students regarded the intervention of CL as a 'very helpful' or 'helpful' approach to learning website architecture and web page programming (see Fig. 1).

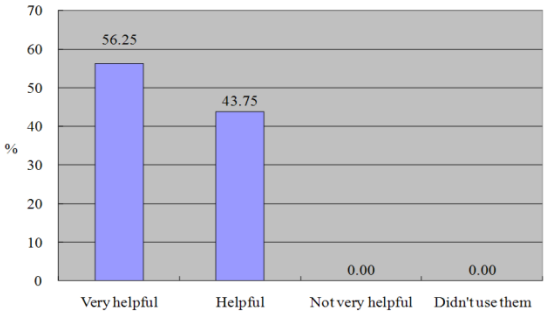


Fig. 1. Helpfulness of CL for developing students' computing skills

Students were asked whether SRL and the mechanism of submitting homework within the required time was helpful for developing regular behavior of learning. As shown in Fig. 2, 48.43% of students thought this was very helpful, with 89.06% regarding the mechanism as 'helpful' or 'very helpful'.

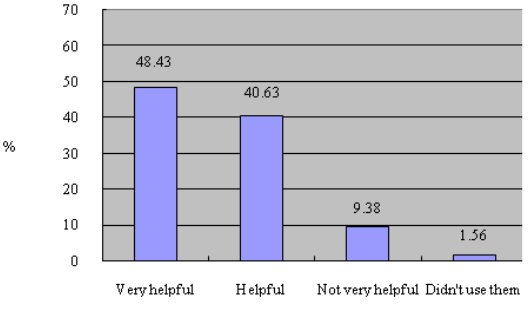


Fig. 2. Helpfulness of SRL for developing regular behavior of learning

Fig. 3. Comparison of learning experiences in taking this blended course with traditional courses

Furthermore, students were also asked to compare the experience of this blended course with traditional courses they had taken before. As shown in Fig. 3, 96.87% judged this course integrating SRL and CL as 'much better' or 'better' than traditional courses that they had taken before.

4 Discussion and Conclusion

There is a world-wide trend in educational policy and research to realize the need to reform education into more effective and innovative forms of pedagogical practice [6]. In this study, the author redesigned a course, adopted web-mediated SRL and CL, and explored their effects on improving students' thoughts regarding this course and the innovative teaching methods. The results of this study indicate the positive effects of SRL, CL, and online learning on improving students' experiences of online learning, and their positive thoughts regarding the interventions concerning SRL and CL. Finally, it is expected that the interventions of web-mediated SRL and CL and the results in this study could provide references for teachers in nations which have delivered didactic pedagogy for many years to adopt innovative teaching methods when providing online or blended courses.

References

1. Lee, T.H., Shen, P.D., Tsai, C.W.: Enhancing computing skills of low-achieving students via e-learning: A design experiment of web-based, problem-based learning and self-regulated learning. CyberPsychology & Behavior 11, 431–436 (2008)
2. Zurita, G., Nussbaum, M.: Computer supported collaborative learning using wirelessly interconnected handheld computers. Computers and Education 42, 289–314 (2004)
3. Cennamo, K.S., Ross, J.D., Rogers, C.S.: Evolution of a web-enhanced course: Incorporating strategies for self-regulation. Educause Quarterly 25, 28–33 (2002)
4. King, F.B., Harner, M., Brown, S.W.: Self-regulatory behavior influences in distance learning. International Journal of Instructional Media 27, 147–156 (2000)
5. Boyle, T., Bradley, C., Chalk, P., Jones, R., Pickard, P.: Using blended learning to improve student success rates in learning to program. Learning, Media and Technology 28, 165–178 (2003)
6. Ottestad, G.: Innovative pedagogical practice with ICT in three Nordic countries – differences and similarities. Journal of Computer Assisted Learning 26, 478–491 (2010)

A Novel Simplified Log-Likelihood Ratio for Soft-Output Demapping of CMMB

Xiangran Sun[1] and Zhibin Zeng[2]

[1] Information Engineering School
[2] Engineering Research Center of Digital Audio & Video, Ministry of Education
Communication University of China
Beijing, China
sunxr@cuc.edu.cn

Abstract. In this paper, a new soft-output demapper generating log-likelihood ratio (LLR) is proposed based on the simplified algorithm instead of an exact LLR. Moreover we design an efficient soft-decision demapping algorithm and implement a low complexity demapper for the 16-QAM constellation of the China Mobile Multimedia Broadcasting (CMMB) standard. In addition, the proposed soft-decision demapper can reduce the hardware complexity by discarding the multipliers and decreasing the soft-value width.

Keywords: demapping, log-likelihood ratio (LLR), quadrature amplitude modulation (QAM).

1 Introduction

China Mobile Multimedia Broadcasting is becoming widespread and received in many city today. As a demand for high quality audios and videos services increase, the broadcasting company needed a wireless digital broadcasting system which is suitable for mobile handout terminals. As a result, the standardization work of CMMB was completed in 2006 [1]. Its performance was improved since the advanced modulation, called orthogonal frequency division multiplexing (OFDM), and the powerful coding schemes, namely the low-density parity-check (LDPC) codes, are utilized.

OFDM has been chosen as the modulation format because of its good performance in highly dispersive channels. The 16 quadrature amplitude modulation (16-QAM) used as mandatory sub-carriers modulation formats of OFDM in CMMB. The receiver can export the hard- or soft-output by demapping the 16-QAM after de-modulating OFDM symbols. In order to achieve the performance required by, a received signal should be soft-decided rather than hard-decided. Therefore, it is needed for a technique of soft-deciding received symbols as bits in higher-order modulation system. LDPC codes were first presented by Gallager [2] in 1962. LDPC codes use log-likelihood ratio (LLR) for probability based decoding such as belief propagation, depending on the soft-output of the demapper [4]. In addition, it achieves the higher performance in terms of available Signal-to-Noise Ratio (SNR).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 350–356, 2011.

The LLR is generated by a demapper using posteriori probabilities, and the demapper is supposed to pass the exact LLR for each bit of codeword to the decoder of LDPC [4]. The exact LLR reflects full information of likelihood functions for all possible constellation symbols. In fact, in the case of the 16-QAM constellations, each axis carries more than one bit and the metric functions, in the soft-output demapper, to determine the soft information for each bit, are in general quite complicated [4].

In the conventional wireless communication system, LLR method has been used as the soft-decision technique. However, this method has problems in terms of hardware complexity and power consumption due to complicated operations [6]. To reduce the high hardware complexity, we propose a low-complexity soft-decision demapper based on the simplified method.

The rest of this paper is organized as follows. Section 2 describes the system model of CMMB. Section 3 presents the soft-output demapping algorithms particularly. Section 4 proposes the new simplified algorithm for 16-QAM of CMMB, and also presents the simulation results. Finally, conclusions are drawn in Section 5.

2 System Model

In Figure 1 the CMMB system model on which we base our analysis is shown. At the transmitter, the input bits, after LDPC encoded with a rate 1/2 and 3/4, bit-by-bit interleaved and then converted into QAM symbols, according to Gray-coded constellation mappings. The complex symbols are then fed to an OFDM modulator, arranged into a physical frame called physical (PHY) burst, and finally transmitted.

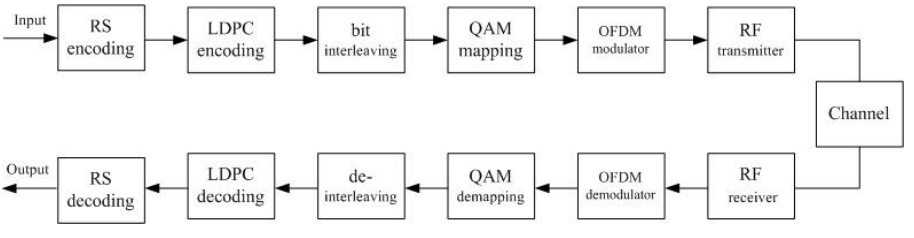


Fig. 1. Block diagram of CMMB system model

Let $M=2^{2m}$ be the number of symbols of the generic square QAM constellation, so that m interleaved bits are mapped into the in-phase and quadrature components of the complex symbol.

Let $a[i]=a_I[i]+ja_Q[i]$ denote the QAM symbol transmitted in the i-th sub-carrier and $\{b_{\mathrm{I},1},b_{\mathrm{I},\mathrm{k}},\cdots,b_{\mathrm{I},\mathrm{m}},b_{\mathrm{Q},1},\cdots,b_{\mathrm{Q},\mathrm{k}},\cdots,b_{\mathrm{Q},\mathrm{m}}\}$ the corresponding bit sequence. Assuming that the cyclic prefix completely eliminates ISI (Inter OFDM Symbol Interference) and ICI (Inter Channel Interference), then the received signal in the generic sub-carrier can be written as

$$r[i]=G_{ch}(i)\bullet a[i]+\omega[i] \tag{1}$$

Where $G_{ch}(i)$ is the Channel Frequency Response (CFR) complex coefficient in the i-th sub-carrier and $\omega[i]$ is the complex Additive White Gaussian Noise (AWGN) with variance $\sigma^2 = N_0$.

At the receiver, the OFDM demodulator performs synchronization and channel estimation, the latter being used by a zero-forcing equalizer which compensates for attenuation and phase shift in each data sub-carrier. If the channel estimate is error free, the output of the one-tap equalizer is given by

$$y[i] = a[i] + \omega[i] / G_{ch}(i) = a[i] + \omega'[i] \tag{2}$$

Where $\omega'[i]$ is still complex AWGN noise with variance $\sigma'^2(i) = \sigma^2 / \left|G_{ch}[i]\right|^2$.

3 Conventional Soft-Output Demapping with Exact LLR

In generally the received QAM signals are first demodulated by a soft-output demapper and de-interleaved, and then passed to a binary soft-input LDPC decoder. The idea is to demap the received signal into soft bits which have the same sign as provided by a hard detector and whose absolute value indicates the reliability of the decision.

For each symbol $r[i]$, 4m metrics need to be derived, two for each in-phase and quadrature bit $b_{I,k}, b_{Q,k}$, corresponding to possible values 0, 1. The optimum hard decision on bit $b_{I,k}$ (the same applies for bit $b_{Q,k}$) is given by the rule $\hat{b}_{I,k} = \beta$ if

$$P\left[b_{I,k} = \beta \middle| r[i]\right] > P\left[b_{I,k} = (1-\beta) \middle| r[i]\right], \beta = 0,1 \tag{3}$$

Set $\beta = 1$, then (3) can be rewritten as $\hat{b}_{I,k} = 1$ if

$$\log \frac{P\left[b_{I,k} = 1 \middle| r[i]\right]}{P\left[b_{I,k} = 0 \middle| r[i]\right]} > 0 \tag{4}$$

Thus, the Log-Likelihood Ratio of decision $\hat{b}_{I,k}$ is defined as

$$\begin{aligned} LLR(b_{I,k}) &\triangleq \log \frac{P\left[b_{I,k} = 1 \middle| r[i]\right]}{P\left[b_{I,k} = 0 \middle| r[i]\right]} \\ &= \log \frac{\sum_{\alpha \in S_{I,k}^{(1)}} P\left[a[i] = \alpha \middle| r[i]\right]}{\sum_{\alpha \in S_{I,k}^{(0)}} P\left[a[i] = \alpha \middle| r[i]\right]} \end{aligned} \tag{5}$$

which is the soft bit information assigned to bit $b_{I,k}$, namely $S_{I,k}^{(0)}$ comprising the symbols with a '0' in position (I, k) and $S_{I,k}^{(1)}$ which is complementary.

By applying Bayes rule and assuming that the transmitted symbols are equally distributed, relation (5) yields

$$LLR(b_{I,k}) = \log \frac{\sum_{\alpha \in S_{I,k}^{(1)}} p\left(r[i] \middle| a[i] = \alpha\right)}{\sum_{\alpha \in S_{I,k}^{(0)}} p\left(r[i] \middle| a[i] = \alpha\right)} \tag{6}$$

Sub-optimal simplified LLR can be obtained by the log-sum approximation: $\log \sum_j z_j \approx \max_j \log z_j$, which is good as long as the sum in the left-hand side is dominated by the largest term, as typically occurs in channels with high Signal-to-Noise Ratio (SNR). Thus

$$LLR(b_{I,k}) \approx \log \frac{\max_{\alpha \in S_{I,k}^{(1)}} p\left(r[i] \middle| a[i] = \alpha\right)}{\max_{\alpha \in S_{I,k}^{(0)}} p\left(r[i] \middle| a[i] = \alpha\right)} \tag{7}$$

Since the conditional pdf of $r[i]$ is complex Gaussian

$$p\left(r[i] \middle| a[i] = \alpha\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2} \frac{\left|r[i] - G_{ch}(i)\alpha\right|^2}{\sigma^2}\right\} \tag{8}$$

Combine (8) and (7) and normalizing by $2/\sigma^2$ the final soft bit values can be calculated as:

$$\begin{aligned} LLR(b_{I,k}) &= \frac{\left|G_{ch}[i]\right|^2}{4} \left\{ \min_{\alpha \in S_{I,k}^{(0)}} \left|y[i] - \alpha\right|^2 - \min_{\alpha \in S_{I,k}^{(1)}} \left|y[i] - \alpha\right|^2 \right\} \\ &= \left[m_0(b_{I,k}) - m_1(b_{I,k}) \right] / 4 \end{aligned} \tag{9}$$

Where defined $m_0(b_{I,k}) = \min_{\alpha \in S_{I,k}^{(0)}} \left|y[i] - \alpha\right|^2$ and $m_1(b_{I,k}) = \min_{\alpha \in S_{I,k}^{(1)}} \left|y[i] - \alpha\right|^2$. Yet, expression (9) allows further simplification with significant reduction of computational complexity and negligible soft-decoding performance loss.

4 Simplified LLR for CMMB

4.1 Simplified LLR Algorithm

Figure 2 shows the 16-QAM mapping scheme in CMMB. Every point in the constellation is mapped to four bits (b_3, b_2, b_1, b_0). As we can see the (b_3, b_1) corresponding to $(b_{I,1}, b_{I,2})$, and (b_2, b_0) corresponding to $(b_{Q,1}, b_{Q,2})$ they are delimited by either horizontal or vertical boundaries. Therefore, the two symbols within the two subsets, nearest to the received equalized signal, always lie in the same row if the partition boundaries are vertical or in the same column if the boundaries are horizontal. As a consequence, equation (9) can be rewritten as:

$$\begin{aligned} LLR(b_{I,k}) &= \left|G_{ch}[i]\right|^2 \bullet \left\{ \left(y_I[i] - \alpha_{I,0} \right)^2 - \left(y_I[i] - \alpha_{I,1} \right)^2 \right\} \\ &= \left|G_{ch}[i]\right|^2 \bullet \left\{ \left(\alpha_{I,1} - \alpha_{I,0} \right) \left[2y_I[i] - \left(\alpha_{I,1} + \alpha_{I,0} \right) \right]^2 \right\} \end{aligned} \tag{10}$$

Evaluation of the terms $D_{I,k}$ for the in-phase bits of a 16-QAM symbol yields

$$D_{I,k} = \left(\alpha_{I,1} - \alpha_{I,0} \right) \left[2y_I[i] - \left(\alpha_{I,1} + \alpha_{I,0} \right) \right] \tag{11}$$

It can be easily verified that the $D_{Q,k}$ functions for the two quadrature bits are the same as (11) with $y_I[i]$ replaced by $y_Q[i]$.

Equation (11) is very complex to implement on hardware, so we introduce a further simplification. It can be approximated as follows:

$$D_{I,1} \simeq -y_I[i] \tag{12}$$

$$D_{I,2} \simeq \left| y_I[i] \right| - 2/\sqrt{10} \tag{13}$$

$$D_{Q,1} \simeq -y_Q[i] \tag{14}$$

$$D_{Q,2} \simeq \left| y_Q[i] \right| - 2/\sqrt{10} \tag{15}$$

It should be emphasized that since the derived expressions for the LLRs in (12) to (15) are not derived with any particular decoding algorithm in mind, they can be used with any iterative decoder such as LDPC and turbo decoders. These approximated expressions still only require a fraction of the complexity as compared to (6). If we set the noise variance σ^2 in these expressions to an arbitrary scalar, as commonly done when using these decoders, we obtain the same expressions in (12) to (15).

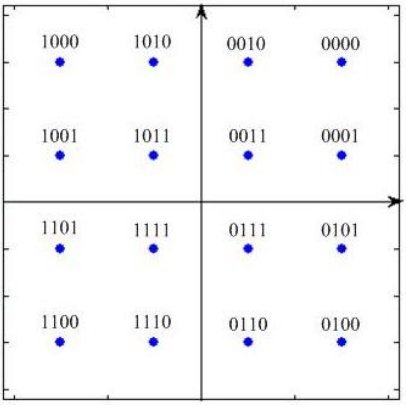


Fig. 2. 16-QAM mapping scheme of CMMB

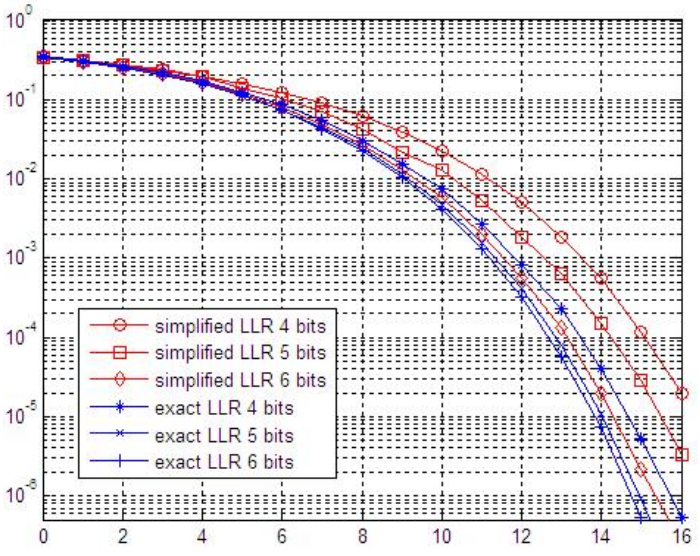


Fig. 3. Bit error performance comparison between exact and simplified LLR demappers of soft bit values for 4, 5 and 6 bits width with 16-QAM of CMMB on AWGN channels

4.2 Simulation Results

The results, obtained by computer simulations, are given in terms of Bit Error Rate (BER) versus E_b/N_0, where E_b is the energy per information bit. In order to compare the performance of the exact and simplified LLR demappers for the three soft bit values which can be used, we select 4, 5, 6 bits width of soft bit value to simulate. The output soft bit value is very important to the performance and the hardware consumption. If the soft bit value is smaller, the hardware consumption is lesser, but the performance is lower. We get the balance of the performance and consumption from our algorithm simulation. In each case adequate soft-output bit demappers have been implemented in exact and simplified forms on AWGN channels.

Figure 3 shows the bit error performance for the 16-QAM of CMMB as a function of the SNR per bit under different conditions: exact and simplified LLR, 4, 5 and 6 bits width of soft bit value. As can be observed, there is a substantial difference in the performance level between exact and simplified LLR. Figure 3 also shows that there is a very good similarity between the exact and simplified LLR algorithms, when the width of soft bit value is 6 bits. Considering the hardware consumption the 5 bits soft-output demapper with simplified LLR is the optimal scheme for CMMB.

5 Conclusion

In this paper a simplified LLR algorithm and a low-complexity soft-output demapper for applications in 16-QAM of CMMB have been designed and developed. Basically, this has been accomplished by simplifying the LLR in each case and using the small soft bit value. Whole design has been applied to the implementation of the demapper

for CMMB. Ongoing and future work includes the improvement of algorithm performance and the reduction of the hardware resource.

References

1. Mobile Multimedia Broadcasting (P. R. China) Part 1: Framing Structure, Channel Coding and Modulation for Broadcasting Channel
2. Gallager, R.G.: Low-Density Parity-Check Codes. IEEE Trans. on Inform. Theory 8, 21–28 (1962)
3. Hagenauer, J., Oer, E., Papke, L.: Iterative decoding of binary block and convolutional codes. IEEE Transactions on Information Theory 42(2), 429–445 (1996)
4. Lee, I., Chang, D., Oh, D.: Multi-Level Modulation LDPC Decoding Algorithm for New Generation DVB-S2 System. In: Proc. 24th AIAA
5. Park, J.W., Sunwoo, M.H., Kim, P.S., Chang, D.-I.: Low complexity soft-decision demapper for high order modulation of DVB-S2 system. In: ISOCC 2008, International, vol. II, pp. 37–40 (November 2008)
6. Tosato, F., Bisaglia, P.: Simplified Soft-Output Demapper for Binary Interleaved COFDM with Application to HIPERLAN/2. In: IEEE International Conference on Communications, ICC 2002, vol. 2, pp. 664–668 (August 2002)

The Design of the Open Digital-Classroom System and Its Workflow

Juhong Xu[1], Di Chen[2,*], and Lin Li[2]

[1] School of Art& Design, Wuhan Institute of Technology, Wuhan, China
[2] Department of Information Technology, Central China Normal University, Wuhan, China
x2u2@sohu.com,
{chendi,linli}@mail.ccnu.edu.cn

Abstract. The information technology based on computers is developing rapidly, which provides basis and possibilities for digital learning. So it is very meaningful to develop a digital learning environment to support present learning with the help of network technology, virtual reality technology and multimedia technology. In the basis of analyzing the users of instruction system, this paper has discussed detaily the design of the open digital-classroom system and its workflow. And the goal is that these studies are helpful to build a new completed digital environment for E-Education.

Keywords: digital-classroom system; design; workflow; classroom teaching; network learning.

1 Introduction

With the development of information age, education and its teaching models are developing and changing constantly. At present, the school-based classroom teaching system and the web-based distance learning system are two most important forms. But both of them cannot satisfy nowadays' education, So in the current digital age, blending the advantages of the traditional instruction and network learning and discarding their disadvantages to develop a digital learning environment that adapt to nowdays' learning is very meaningful.

2 The Analysis of System Users

Instruction system is the environment designed for implementation of teaching activities. For the school-based teaching system, the users are local teachers and local students; for the web-based distance learning system, the users are divided into local teachers, local students and network students.

With regard to the user of teachers, they need diverse show channel to show course contents, they need supports to design teaching plans and implement diverse teaching activities, and they need real-time and intuitional data to evaluate and

* Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 357–363, 2011.

feedback in teaching process, moreover, the instruction media must be convenient and stable to use for them.

With regard to the user of students, not only they need to watch and listen when participating in teaching activities, but also need to think and discuss questions raised by teachers and fully express their views; further they need to complete theme activities around study topic, find and read the relevant information through the supports of technology, and at last they earn their own understands. Especially for network students, how to participate in classroom activities, discuss with others and express their own views and so on, all these demands need supports of instruction system.

In addition, the users of system include management and maintenance crew. Their responsibilities are to manage and maintain daily hardware and software in instruction systems, and record, organize and publish teaching process resources.

3 The System Function Model, the System Hierarchical Structure Model And the System Hardware Topological Model

The system function model of the open digital-classroom system can be designed as the following figure:

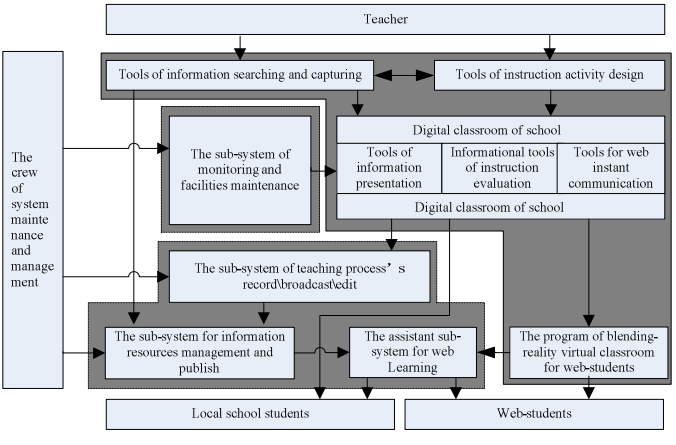


Fig. 1. The system function model

The system function structure can be divided into three modules: the classroom teaching tool module, the network teaching tools module and the system management and maintenance module.

*1)The classroom teaching tool module.*On the one hand this module assists teachers to complete plan design of informationzation instruction activities, on the

other hand it provides real-time classroom teaching environment to local students and network students.

*2)The network teaching tool module.*This module includes sub-system for information resources management and publish, the assistant sub-system for web learning and the sub-system of teaching process's record\broadcast\edit. This part records, manages and publish teaching process.

*3)The system management and maintenance module.*It includes the network monitoring and maintenance sub-system. The three functions that management and maintenance crew need can achieve through it.

Figure 2 is the system hierarchical structure model:

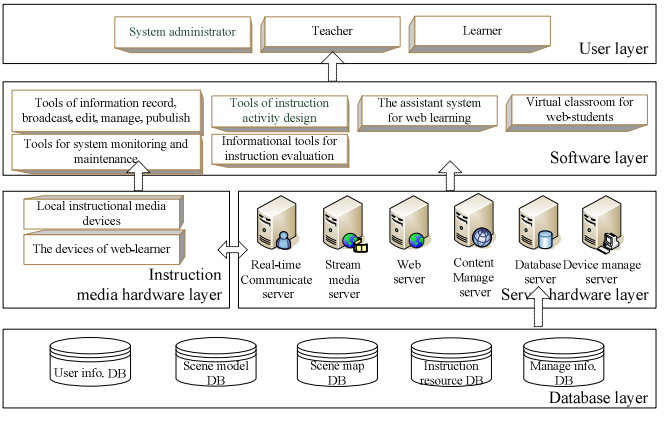


Fig. 2. The system hierarchical structure model

The whole system can be divided into four layers: database layer, hardware layer, software layer and user layer.

Database layer includes user information database, scene model database, scene map database, instruction resource database and manage information database, all these organize and manage information of the system.

Hardware layer includes server hardware layer and instruction media hardware layer. Server hardware layer saves and distributes network interaction information in teaching process. Instruction media hardware layer is to provide direct information display, reduction, communication and other hardware supports.

Software layer includes tools of information record\broadcast\edit\manage\publish, tools of instruction activity design, tools for system monitoring and management, information tools for instruction evaluation, the assistant system for web learning, virtual classroom for web-students, and so on.

The user layer includes administrators, teachers and students.

Figure 3 is the system hardware topological model:

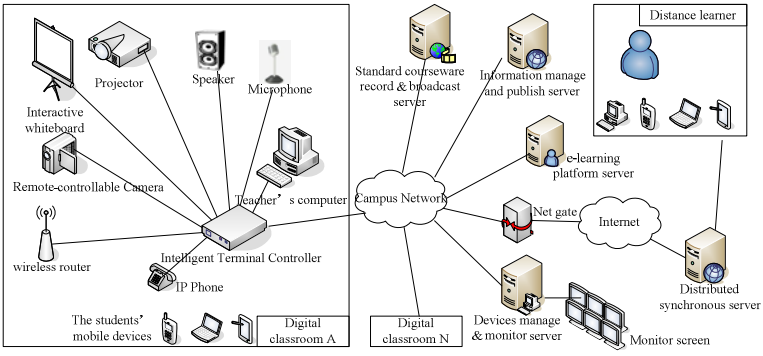


Fig. 3. The system hardware topological model

4 The System Workflow

4.1 The System Workflow

The workflow of the open digital-classroom system is that:

First the teacher carries through course design or teaching design before class, the resources that searched and processed in design processing are uploaded to resources service management platform for learners to use (preview or review), and secondly courseware resources can be directly used in digital classroom for classroom teaching. The classroom teaching process is facing local students, and it also can be maked into network video and audio resources that facing network students through operation crew's processing.

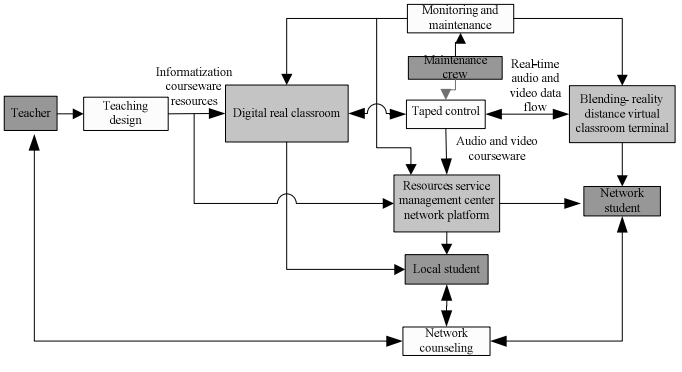


Fig. 4. The workflow of the open digital-classroom system

Network video and audio resources has two purposes, one is to send to remote interactive terminals through real-time broadcast transmission ,and to supply network students to participate in classroom teaching real-timely; Meanwhile, the video and audio infromation of the network learners are transmitted to digital classroom directly, so the teacher can interact with the students. On the other hand the courseware generation tool can make this data into teaching courseware that published in resources

service management platform, and which can be shared by students for asynchronous autonomous learning afer class.

In the system ,the students not only can participate in the classroom teaching real-timely, but also can raise questions and seek help through the tutorial and answering questions system of the network service management platform, and the other students and teachers will give help and advices. Also the teacher can give some studing resources or advices before or after class to learners,let them do preview and review work well.

This process requires management crew to process all kinds of requests and fault early warning of the hardware, inspect system hardware operating condition and the safety of website platform regularly, etc.

4.2 The Teacher's Teaching Workflow

The teacher's teaching process can be divided into three stages.

The first is the teaching design stage, the teacher design teaching activities according to the teaching goal, content, object, environment and such on aspects, take out teaching activities plans and corresponding resources.

The second stage is real-time teaching stages, the teacher teaches in the open digital classroom, and he uses informationization tools for in-class teaching evaluation, revises teaching segments or process and assign homework according to the evaluation results.

The third is the after-school consolidated stage, on the one hand, the teacher judges whether the teaching goal is obtained or not according to homework,if not, then he gives learning guidance through network aided instruction platform; On the other hand the teacher must answer the students'questions. The process is shown below:

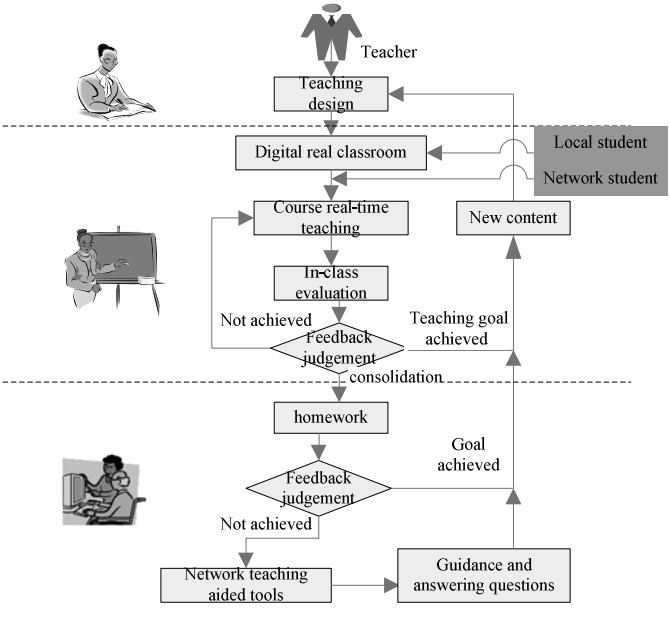


Fig. 5. The teacher's teaching process model

4.3 The Student's Learning Workflow

The student's teaching process is divided into two stages.

The first is real-time classroom learning stage, for local students ,they study in the real digital classroom; for network students,they study in network virtual classroom in form of immersive real-time classroom learning.their quesitions can answered in real-time classroom through network commnication tools.

The second is after-class learning stage,the students can solve their difficulties in studying through browsing the network information (including the classroom teaching process video) ,if they can not solve them,they can seek teacher's help through asynchronous network communication platform. The process is shown below:

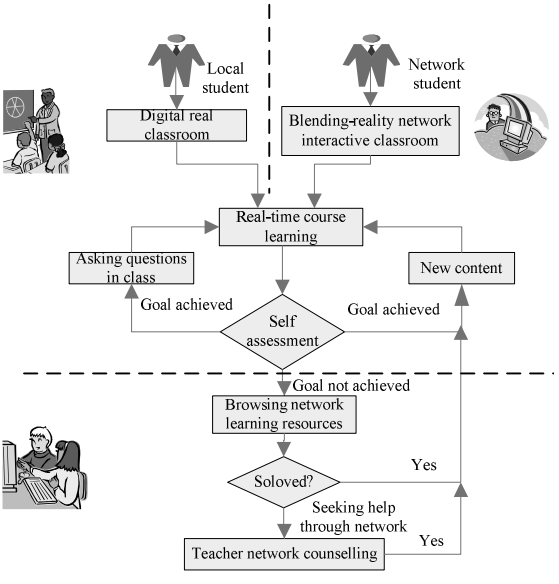


Fig. 6. The student's teaching process model

5 Conclusion

The studies of the design of the open digital-classroom system and its workflow are in order to build a new completed digital environment for E-Education, further promote the development of digital learning situation ,expand E-Education deeply, and boost network learning and traditional classroom teaching mixing together.

References

1. Savidis, A., Grammenos, D., Stephanidis, C.: Developing inclusive e-learning systems. Universal Access in the Information Society (1), 51–72 (2006)
2. Hoyles, C., Noss, R., Kent, P.: On the Integration of Digital Technologies into Mathematics Classrooms. International Journal of Computers for Mathematical Learning (3), 309–326 (2004)

3. Molka-Danielsen, J., Carter, B.W., Richardson, D., Jaeger, B.: Teaching and learning affectively within a virtual campus. International Journal of Networking and Virtual Organisations 6(5), 476–498 (2009)
4. Hayashi, T., Watanabe, K., Hayashida, Y., Kodo, H.: Remote Lecture Based on Instruction with Blackboard Using High-Quality Media Systems. In: Proceedings of the International Conference on Computers in Education, ICCE 2002, p. 858 (2002)
5. Wilson, B.G.: Designing E-learning environments for flexible activity and instruction. Educational Technology Research and Development (4), 77–84 (2004)

Reconstruction and Representation for 3D Implicit Surfaces

Chung-Shing Wang[1,*], Teng-Ruey Chang[2], and Man-Ching Lin[3]

[1] Department of Industrial Design, P.O. Box 965, Tung-Hai University, Taichung, Taiwan
Tel.: 886-4-23590492 ext. 26, Fax: 886-4-23594645
cswang@thu.edu.tw
[2] Department of Industrial Engineering and Management,
Nan-Kai University of Technology, Nan-Tou, Taiwan
[3] Department of Industrial Engineering and Systems Management,
Fung-Chia University, Tahchung, Taiwan

Abstract. Radial Basis Function (RBF) Kernel method is currently the most useful method for carrying out 3D implicit surface reconstruction. However, fitting RBF to 3D scattered data has not been regarded as computationally feasible for large data sets. For this reason, this research conducts an in-depth investigation on implicit surface construction, along with self organizing map (SOM) network and kernel method both in theory and experiment. From research results, we can then use SOM network to obtain geometric features which describe the original model. The use of kernel methods makes calculation of the implicit surface more simple and efficient for performing broken surface reconstruction.

Keywords: Reverse Engineering, Surface Reconstruction, Implicit Surface, Neural Network, Kernel Methods, Self Organizing map.

1 Introduction

3D curve surface construction technique has a wide range of applications such as: industrial design, engineering analysis, virtual reality, animation, medical imaging…etc. Although different applications have different focus, all of them ultimately strive to achieve the optimum geometric reconstruction of the real world. For 3D solid geometric modelling, parametric surface and implicit surface are the two most important 3D surface construction techniques. Normal implicit surface construction uses Radial Basis Function (RBF) kernel method along with structural interpolation calculation to find out coordinates of any point on the curve surface (including coordinates of missing positions) to give an output of any resolution we desire. This also ensures that the vectors and slope of any position on the curve surface are resolvable [1, 2, 3]. However, implicit surface reconstruction technique still has some limitations. For example, large point clouds data not only require a

* Corresponding author.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 364–372, 2011.

large amount of system resources to store the matrix information, it also means that a large number of RBF centers will create a heavy burden on system in terms of calculation speed thus limiting the effectiveness of implicit curve construction based on RBF method.

Due to the problem mentioned above, this research suggests an alternative method to reduce the number of RBF centers [4]. The original center appendage and accuracy determination procedures are replaced by a Self Organizing Map (SOM) network for 3D point group pre-processing to extract geometric features from object geometries. This results in a reduction in the amount of data necessary to construct the surface. With the use of SOM network, the system can automatically extract the point clouds feature data required for reconstruction of the original geometric functions. Finally, the kernelization algorithm is used to come up with a reasonable function as basis of approximation to increase efficiency of reconstruction with implicit surface technique.

This research investigated the point clouds feature extraction mechanism with the use of SOM network as basis for research development. SOM belonged to an area of Artificial Neural Network (ANN), similar to non-supervised learning in machine learning. In the research on implicit surface, the kernelization algorithm provided us with a reasonable approximation function as basis for further calculation. With the use of dimensional transformation and the concept of vector interpolation, we can improve the simplicity and speed of implicit surface formula construction. In order to solve the problem of non-linearity that had arisen, this research used the Matlab program for simulation of theory and experimentation.

In order to solve the problems of the large amount of system resource required for implicit curve calculation and also incorrect reconstruction due to local divergence, this paper suggested the use of a partial calculation method as opposed to the structural tree method [5]. The SOM calculation method of ANN was used to obtain the geometric topology with reference to the original model to carry out global feature mapping of point clouds data. Corresponding feature points were then created and finally, with these points as kernel centers, the implicit surface was calculated.

Extraction of feature points through SOM network, with the purpose of retaining feature characteristics in the original model, can significantly reduce the amount of data that the image required. This meant that there was greater efficiency when carrying out implicit curve construction in later steps. Interpolation was then used to create the implicit surface in order to retain the completeness of model geometry so that further model data calculation can proceed more smoothly.

A representation of feature extraction and kernel surface reconstruction of incomplete point clouds was shown in Fig. 1. It can be separated into the following five steps:

(1) 3D Scan Data: Reverse scanning engineering was used to obtain data of model surface coordinates.
(2) Network Initialization: SOM network structure was defined along with setting learning constant and learning termination criteria.
(3) Iteration: Winning neuron was calculated and weight vector correction carried out. Iteration was repeated until termination criteria were satisfied.

(4) Geometric Features: Learning method of SOM network was used to carry out feature mapping of point clouds. Feature points were then extracted from the model. Finally, geometric features of the model were created with Delaunay triangulation.

(5) Implicit Surface: Kernel method was used to carry out 3D structure vector interpolation of geometric features to create corresponding implicit surface equations. Zero level iso-surface calculations were then carried out with the use of meshes to represent surface geometry.

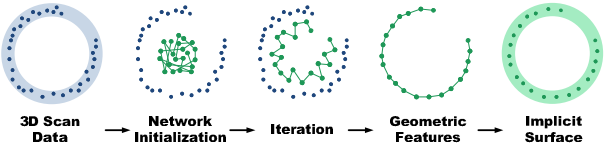


Fig. 1. Reconstruction for implicit surface

2 Paper Construction

The objective of this research was to use SOM network and RBF kernel method to carry out implicit surface reconstruction. Coordinate measuring machine (CMM) was used to obtain point clouds data; feature outlines were then extracted; finally, corresponding kernel function estimations were created. The research flow diagram was shown in Fig. 2.

(1) Feature extraction: Non-supervised SOM feature mapping was used to extract outline geometry from the original model. Geometric feature points that can replace the original structure were selected to complete feature extraction.

(2) Surface reconstruction: RBF kernel calculation with respect to feature point clouds data was used to carry out vector interpolation for working out the implicit surface function.

2.1 Feature Extraction

The 3D point clouds data obtained from digital measurement in reverse engineering can be sent to SOM network for self-competitive learning to extract implicit feature geometries. Delaunay Triangulation was then used to carry out model pavement to recreate feature geometries in order to reduce the memory space required for the system to store point clouds data. This served to minimize the large amount of system calculation required for computing implicit surfaces.

SOM, also referred to as the Kohonen Map, was mentioned by Kohonen as a special example of non-supervised learning algorithm belonging to the ANN competitive learning rule [6]. The SOM operation was shown in Fig. 3. Vectors of observation samples were first entered and projected onto the corresponding neuron with feature mapping. Feature mapping was then used as a non-linear projective model to convert the input vectors into matrix space composed of neurons. From competition between vectors, the winning neuron was determined and used as the base weight

vector for correction. The neuron of the final output layer (or Kohonen layer) was displayed in the output space according to corresponding connections within the network topology in relation to features of the input vectors. Because the topology structure reflected the features of the observed sample, the whole process mimicked the perception mapping of the brain and therefore was named SOM.

The previously mentioned SOM Feature Extraction flowchart (Fig. 4) can be separated into five main steps:

(1) Initialize the SOM network: First of all, the number of Neurons and learning parameters were set. A set of random numbers between 0~1 were then initialized corresponding to the Neuron vector values.

(2) Select the best matching cell: Inputted vector value X was used to start the SOM network with p as iteration number. The smallest Euclidian distance between all input vectors and weight vectors were calculated to obtain the winning neuron.

(3) Adapt the weight vectors: Mexican hat function and learning constant with reference to winning neuron were used to define the creation and control of neighboring neurons in order to adjust vector weights [7].

(4) Reach termination conditions: the iteration number was increased by 1 and process repeated from step (2) until termination conditions were met. Termination conditions can be criteria such as: the least Euclidean distance, total iteration numbers or checking the feature corresponding process on the network so that there were no significant changes.

(5) Triangulize the surface: With calculated weight vectors as feature coordinates, the Delaunay Triangulation method was used to create triangular meshes to complete construction of feature geometries.

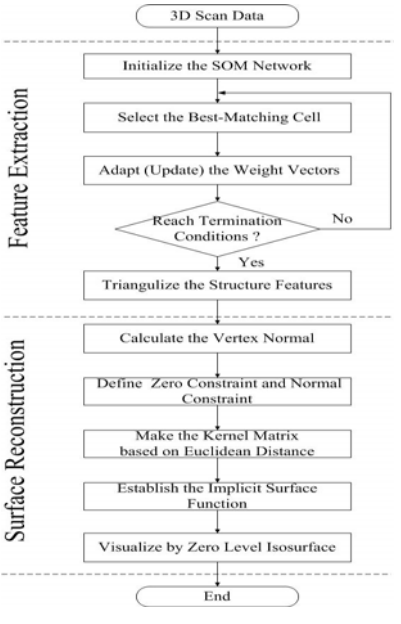


Fig. 2. Process for the research

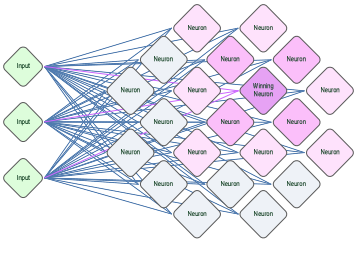


Fig. 3. SOM model based on winning neutron

2.2 Surface Reconstruction

Implicit surface construction can be seen as a signed-distance function problem [8]. This process was comparable to machine learning. However, the kernelization algorithm also provided us a reasonable interpolation calculation basis. With dimensional transformation and vector interpolation theories we can obtain vector distances of all positions in the space. This served as the starting point for implicit surface calculations.

An implicit surface was defined by an implicit equation within continuous scalar function within a 3D space so that when points on the surface were substituted into the function, the result will be zero. Implicit function can be considered as point clouds that satisfy the condition that the implicit number is zero.

$$f(x) = 0, \qquad x \in R^3 \tag{1}$$

On the constructed surface, the said points were on the surface; if the function was positive, the points were outside the surface and if the function was negative, the points were inside the surface. When we define the vector distance between points and the surface as *d*, the implicit surface function mentioned above becomes a number of the function.

$$f(x) = d \tag{2}$$

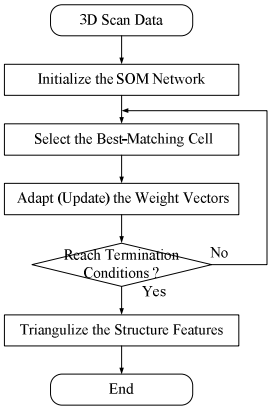


Fig. 4. Process for feature extraction

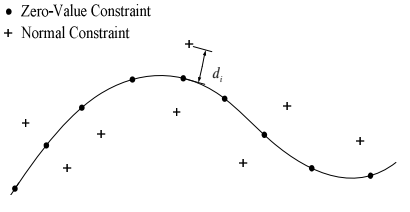


Fig. 5. Signed distance function

Refer to Fig. 5 as an example. This research showed the points on the curve as its zero-value constraints. Points that were not on the surface were considered as surface vector normal constraints for vector distance *d*.

Implicit surface was based on zero-value constraints and normal constraints in addition to function interpolation and zero level iso-surface calculation to recreate surface geometric shape. Detailed calculation flow chart is shown in Fig. 6.

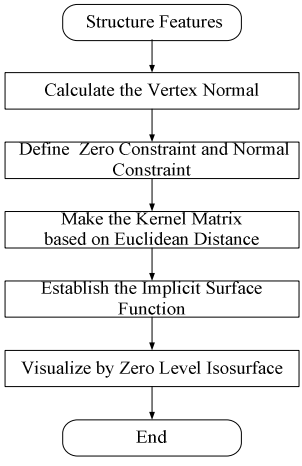


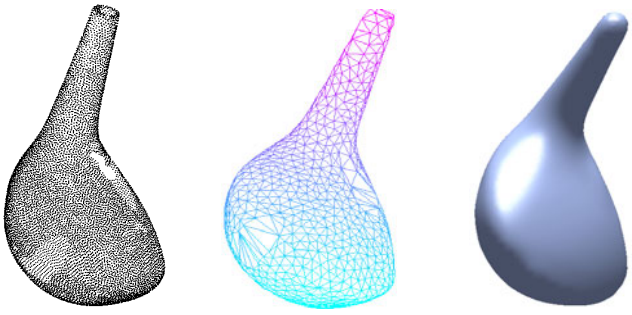
Fig. 6. Kernalization for implicit surface

There are five main steps:

(1) Calculate the vertex normal: First of all, the model feature geometries were consulted. The related triangular surfaces were found from the constructed Delaunay triangulation model. Vector averages were then calculated for these triangular surfaces to become vector values of feature points.
(2) Define zero constraint and normal constraint: Zero constraints were defined by feature point coordinates while Normal constrains were determined by offsetting a small distance from feature point in the vector direction.
(3) Make the kernel matrix based on Euclidean distance: All the zero constraints and normal constraints were considered when calculations of Euclidean distance were carried out. With the use of kernel function, a kernel matrix was created.
(4) Establish the implicit surface function: With the use of kernel matrix, LU decomposition was carried out on the kernel function estimation to resolve the undetermined coefficients to complete construction of implicit surface function.
(5) Visualize by zero level iso-surface: Zero level iso-surface calculations were carried out on the implicit surface. Implicit function model were extracted with geometric topology in accordance to Marching Cube method. Zero level iso-surfaces were then displayed by triangular meshes.

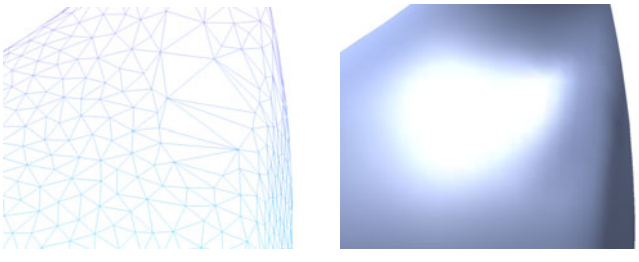
3 Implementation

Current research focused on the use of SOM network to extract point clouds data with kernel method in order to reduce the burden on system when data were accessed and calculated when implicit surfaces were created. For the case study, Chitai LSH 800 3D coordinate measuring machine was used to carry out verification of Golf club model reconstruction. Because noise processing was not part of this research, the noise was pre-processed. To reduce the excess number of data points that caused surface



(a) (b) (c)

Fig. 7. (a)Points data in golf head (No. of points = 10,034) (b) Geometric features for golf model (feature points = 1,024) (c) Implicit surface reconstruction for golf head (interpolation points = 125,000)



(a) (b)

Fig. 8. (a)Feature extraction for broken surface in golf head (b)Implicit surface reconstruction for broken surface

turbulent phenomena, this research also carried out appropriate data reduction to the model. The resulting data was outputted in ASCII format, saved as TXT files and the surface reconstructed with Matlab.

For the extraction of model features, this research used the highest iteration number of the network as termination condition. Under most circumstances, the greater the number of iterations, the higher the quality of extracted features, but this would also result in an overly long calculation time. In the case study with the Golf club model, after multiple validations, it was discovered that iteration number set to 100 times was sufficient to extract enough detail from the network. In terms of structural interpolation, in order to obtain a better visualization effect, this research set the interpolation points at 125,000 points. Marching cube method was used to find out zero level iso-surface and then finally, the implicit surface structure was displayed with triangular meshes.

Multiple scans were done on the actual golf head model to obtain outline data from different angles. These data were then imported into Geomagic Studio software. With the use of point clouds data with merge calculation; a more complete collection of

point clouds data were created. The final number of data points were 10,034 points, see Fig. 7(a). Fig. 7(b) and Fig. 7(c) showed the results of feature extraction and surface reconstruction with a total of 1,024 feature points with 125,000 interpolation points. To calculate the errors of the reconstruction, this research inputted the original data into the implicit surface function to obtain the vector offset values. Those values were used as the errors of reconstruction to acquire root mean square (RMS) error and maximum error. The broken surface of the model was referred to in Fig. 8(a); the result of feature extraction and reconstruction was illustrated in Fig. 8(b). Vector offset values for reconstructed surfaces with different feature points were displayed in Table 1. From the research results, the methods described in this paper can be successfully applied to incomplete data with broken surface geometry to carry out reconstruction of model. We were also able to achieve a significantly positive structural interpolation effect.

Table 1. Errors in normal vector for golf head model

Feature	RMS error(mm)	Max error(mm)	Time(sec)
16	8.727682e-001	2.671489e+000	13.2
32	7.090761e-001	2.621135e+000	18.4
64	2.209513e+000	3.174602e+000	24.0
128	8.790169e-002	4.933744e-001	45.6
256	4.959841e-002	3.201560e-001	83.3
512	3.035979e-002	3.830089e-001	171.1
1024	1.981223e-002	4.226490e-001	403.4

4 Conclusion

This research focused on the use of SOM network and kernel method to extract feature geometry and reconstruct structure though interpolation from incomplete 3D point clouds. Research results showed that SOM network can act as a good data simplification process in addition to the use of non-supervised learning method to carry out 3D point clouds feature mapping. This allowed us to find feature point clouds that represented the original model data for further calculation for construction of implicit surfaces. With application of kernelization algorithm in vector interpolation, calculation for implicit surfaces became more simple and efficient. Structure interpolation effect was also achieved as a result. The feature extraction theory and implicit surface reconstruction technique mentioned in this research can effectively reduce the amount of data for point clouds which led to an increase in file transfer and transportation efficiency. Other applications include 3D surface model for smoothness of water can also be constructed for later engineering researchers for further calculation and analysis.

Acknowledgements

This work is partly supported by National Science Council, Taiwan under the contract NSC 99-2221-E-029-026.

References

1. Savchenko, V.V., Pasko, A.A., Okunev, O.G., Kunii, T.L.: Function respresentation of solids reconstructed from scattered surface points and contours. Computer Graphics Forum 14(4) (1995)
2. Carr, J.C., Fright, W.R., Beatson, R.K.: Surface interpolation with radial basis functions for medical imaging. IEEE Trans. Medical Imaging 16(1), 96–107 (1997)
3. Turk, G., O'Brien, J.: Shape transformation using variational implicit functions. In: SIGGRAPH 1999, pp. 335–342 (1999)
4. Carr, J.C., Beatson, R.K., Cherrie, J.B., Mitchell, T.J., Fright, W.R., McCallum, B.C., et al.: Reconstruction and representation of 3D objects with radial basis functions. In: ACM SIGGRAPH 2000, pp. 12–17 (2001)
5. Holdstein, Y., Fischer, A.: Three-dimensional surface reconstruction using meshing growing neural gas (MGNG). Visual Comput. (24), 295–302 (2008)
6. Kohonen, T.: The self-organizing map. Proc. of IEEE 9(78), 1464–1480 (1990)
7. Berg, M., Cheong, O., Kreveld, M., Overmars, M.: Computational Geometry: Algorithms and Applications, 3rd edn. Springer, Heidelberg (2008)
8. Turk, G., O'Brien, J.: Modelling with implicit surfaces that interpolate. ACM Transactions on Graphics 21(4), 855–973 (2002)

Shape-Based Level Set Method for Breast Mass Segmentation on Breast MRI

Chieh-Ling Huang[1] and Li-Yan Kuo[2]

[1] Dept. of Computer Science and Information Engineering, Chang Jung Christian University
[2] Dept. of Information Management, Chang Jung Christian University
kaio@mail.cjcu.edu.tw

Abstract. This paper proposed a level set method with shape model to segment breast mass on the breast magnetic resonance imaging (MRI). Because the level set model proposed by Chan and Vese can not work well on the breast mass segmentation, this paper adds the shape knowledge into the segmentation method. We will first apply the Chan-Vese level set model to get pre-segmented breast mass and then the position and the size of the pre-segmented breast mass will be calculated to establish the initial shape model. And then this paper used dilation processing to calculate the distance to the shape model contour that taken into consideration to update the level set function. Finally the proposed method is applied to segment the breast mass. Our experimental result will show that the breast mass can be correctly segmented by the above mechanism.

Keywords: breast mass; level set method; image segmentation; shape model; magnetic resonance imaging.

1 Introduction

In recent years, computer aided diagnosis has become more and more advanced, and even become a routine clinical work. The medical images contains valuable information on various applications, such as organs recognition and identification, 3D structure reconstruction, cancer detection, tumor localization on treatment planning. Therefore, medical image segmentation has become an important issue on extracting the useful information in various kinds of medical image. The segmentation result is proved to be nearly as good as the one manually segmented by man in the paper [1], indicating that the technique has been more and more mature.

Breast cancer is the most common type of cancer in the females and breast masses generally present as the major symptom of breast cancer. In order to assist radiologists in detecting masses in the early stage of breast cancer, it is highly desirable to develop a reliable computer aided diagnostic system as an assistant. Breast magnetic resonance imaging (MRI) is a mature technique that is ready for broad clinical use. Therefore, the purposed of this paper is proposes a image segmentation method to segment breast mass on the breast MRI.

There have been many researches on image segmentation topic, and many methods have been proposed such as zero crossing [2], thresholding [3], region based segmentation [4], watershed [5], and level set method [6]. Some of the above methods

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 373–379, 2011.

are gradient-based and are vulnerable to weak edges. Some methods are intensity-based which are vulnerable to noises. Using only the local feature is not sufficient so the global features, such as contour length or area, will be taken into consideration.

If we have some prior knowledge about the partition such as shape, we can integrate the prior knowledge into the segmentation process. This kind of concept has been implemented in many researches [7]. When doing the image segmentation, how to express the contour so that we can easily get the global features and represent the prior knowledge is an important problem. The segmentation model proposed by Mumford and Shah incorporated with the level set approach [8] is a good method to solve this problem.

Base on the above considerations, a level set method with shape model for breast mass segmentation on MRI is proposed in this paper which will be introduced in the following section. The remaining of this paper is organized as follows. Section 2 briefly reviews the level set method and its models. Section 3 describes how to establish the shape model and apply the level set method for breast mass segmentation on MRI image. Section 4 gives the experimental results. Finally conclusions are drawn in Section 5.

2 Level Set Method (LSM)

Level set method is a way to denote active contours [9][10]. For a given image u_0, we can create a level set function $\varphi(x,y)$ with the same size of the image u_0 to describe the contour. The contour is defined as the zero level set of the function φ:

$$C = \{(x, y) \mid \varphi(x, y) = 0\} \tag{1}$$

And the inside region and the outside region of the curve are explicitly defined as:

$$\begin{cases} \varphi(x, y) > 0 & \textit{inside the contour} \\ \varphi(x, y) = 0 & \textit{contour} \\ \varphi(x, y) < 0 & \textit{outside the contour} \end{cases} \tag{2}$$

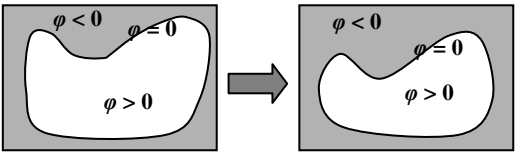


Fig. 1. Illustrations of the level set method and the contour change

By changing the φ values, some regions that are originally negative will turn into positive, and vice versa. Therefore, the contour will change according to the update of the level set function as Fig. 1. There are two main advantages to represent the active contour by level set method. First, it can easily represent complicated contour changes, for example when the contour splits into two or develops holes inside. Second, we can easily know whether a point is inside or outside the contour by checking its φ value.

In this paper, we use the method proposed by Chan and Vese [6] to help us find the contour of the object. The author suggests that the update equation of the φ function should be as

$$\Delta\varphi = \delta(\varphi)\left(\mu \cdot div(\frac{\nabla\varphi}{|\nabla\varphi|}) - \lambda_1(u_0 - c_1)^2 + \lambda_2(u_0 - c_2)^2 - v\right) \tag{3}$$

where u_0 is the gray level value of the original image, c_1 and c_2 are the average gray level values inside and outside the curve, λ_1, λ_2, μ, v are weight parameters, Δt is the step size, and δ_ε is a regularization of the Dirac delta function, which is

$$\delta(\varphi) = \frac{1}{\pi}\frac{\varepsilon}{\varepsilon^2 + \varphi^2} \tag{4}$$

and c_1, c_2 can be easily calculated by

$$c_1 = \frac{\int_\Omega u_0(x,y)H(\phi(x,y))dxdy}{\int_\Omega H(\phi(x,y))dxdy} \quad , \quad c_2 = \frac{\int_\Omega u_0(x,y)(1-H(\phi(x,y)))dxdy}{\int_\Omega (1-H(\phi(x,y)))dxdy} \tag{5}$$

where H is the Heaviside function. A common way to initialize the level set function is by giving a circle as the initial contour and use the signed distance to the circle as the value, which can be calculated by the following equation:

$$\phi_0(x,y) = -\sqrt{(x-c_x)^2 + (y-c_y)^2} + r \tag{6}$$

where (c_x, c_y) and r are the center and the radius of the circle.

In general, the first term of Eq.(3) is the smoothing term which makes the contour smooth and can eliminate some small isolated regions, i.e. noises. The second and third terms of Eq.(3) are the driving force term which drives the contour toward a balanced position. The fourth term of Eq.(3) controls the enclosed area of the level set curve. The Dirac delta function term of Eq.(3) is to scale the update value, which makes the area closer to the contour, i.e. φ=0, to have more significant update than the area farther from contour. We can see that both of the driving force terms do not use any gradient information, but only use the average gray level intensity inside and outside the contour. So this method can work under condition that the edge is not very apparent.

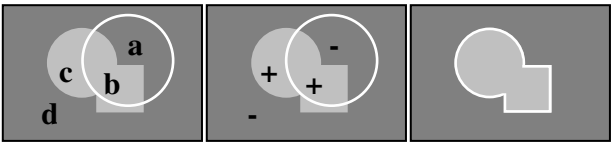


Fig. 2. An example of contour evolution, (a) symbol of each area (b) the sign of update value in each area (c) the evolution result

Now we use an example to explain the evolution of the contour. In Fig. 2, the white circle is the initial contour, so the φ value in **a** and **b** are positive, and the φ value in **c** and **d** are negative. Besides, the gray level intensity in **b** and **c** > c_1 (the average gray level intensity in **a** and **b**) > c_2 (the average gray level intensity in **c** and **d**) > the gray level intensity in **a** and **d**. Hence, the $-(u_0-c_1)^2 +(u_0-c_2)^2$ term is positive in b and c and the term is negative in a and d. After some iterations of update, the φ value in **b** and **c** will become all positive and the φ value in **a** and **d** will become all negative, thus the contour will stop on the desired place.

3 LSM with Shape Model for Breast Mass Segmentation

Figure 3 (a) shows a breast MRI, and Fig.3 (b) shows the result of applying Chan-Vese LSM where the brighter part for the $\varphi>0$ region and the darker part for the $\varphi<0$ region. From result of breast mass part in Fig. 3(c), we can see that there are some noises are also being included. Shape is a good feature for specific image. Therefore the LSM with shape models is the main approach in this paper. In this section, we will describe how to establish the shape model and apply the proposed LSM.

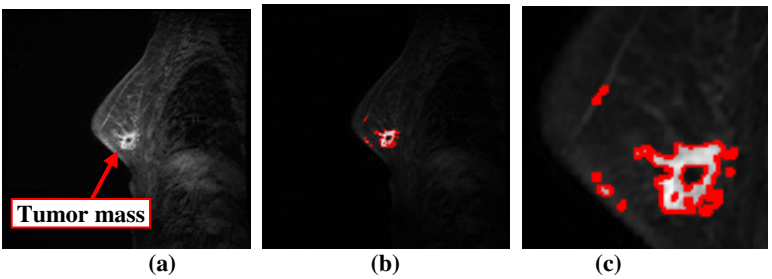


Fig. 3. Result of applying Chan-Vese level set model (a) The breast MRI (b) The result of applying Chan-Vese level set model (c) Part of breast mass

3.1 Establish the Initial Shape Model

To apply level set method with shape prior, we need to establish the initial shape model first. In this paper, we will initialize the shape model manually. Although we know that initializing the shape model automatically is more convenient for image segmentation. However, how to automatically initialize the shape model is depending on the application. There is no initialization algorithm suitable for all cases; therefore we will choose to assign the initial shape model manually. But it is also adequate to use other algorithms to initialize the shape model.

3.2 Resize and Translate Shape Model

Once we get the initial shape mode, we will roughly get the location, and size of the segmented region for the dynamic shape model. We do not need to know the information exactly because we are using the dynamic shape model and the shape model will be adjusted in the segmentation process. The rough information of the segmented region can be got by applying the Chan-Vese level set model using

Eq.(3). And then we can use the location, and the size of the inside region to dynamically adjust the shape model.

3.3 Establish the Distance Map of Shape Model Contour

In our method, the distance to the shape model contour is taken into consideration to update the level set function φ. If the position is far from the contour of the shape model, the level set function φ should be updated rapidly. Therefore, after dynamically adjusting the shape model, we will get the contour of the shape model. And then we will establish the distance map by using the dilation operation to the contour of shape.

Figure 4 illustrates how to establish the distance map by using the dilation operation on the contour of shape. Fig.4 (a) shows the contour of shape, and Fig.4 (b) shows the dilation result of Fig.4 (a). In Fig.4 (b), the values are positive inside and negative outside the region of shape. In the same way, Fig.4 (c)~(e) are the dilated results of dilation of Fig.4 (b)~(d), and the distance increases after next dilation. Finally, we will get the distance map of the shape contour, as shown in Fig.4 (f).

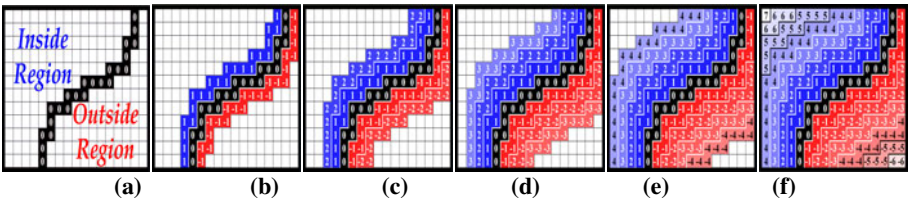


Fig. 4. Establish the distance map by using the dilation operation to the contour of shape.(a) contour of shape (b)~(f) dilation operation to (a) 1, 2,3, 4 and 7 time(s) respectively.

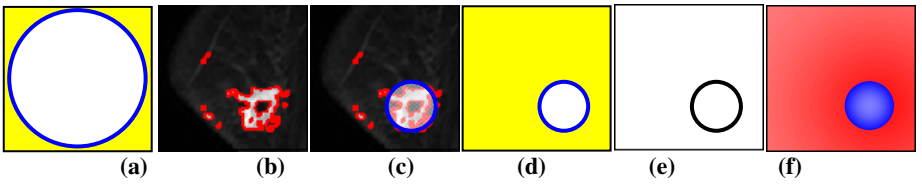


Fig. 5. An example of dynamic adjust shape model, (a) initial shape model (b) level set function φ (c) dynamically adjusted shape model (d) shape model after being dynamically adjusted (e) contour of shape model (f) distance map

Take section 3.1 to section 3.3 for example, Figure 5 (a) shows the manually assigned initial shape model, and Fig.5 (b) shows the level set function φ where the darker region represents the outside region and the brighter region represents the inside region. We will resize and translate the initial shape model according to the level set function φ, as shown in Fig.5 (c). Fig.5 (d) shows the shape model after being dynamically adjusted. After dynamically adjusting the shape model, we will get the contour of the shape model, as shown in Fig.5 (e). Fig.5 (f) shows the distance map where the darker pixel represents shorter distance, the brighter pixel represents

longer distance to the contour of shape and the red region represents the outside region, the blue region represents the inside region.

3.4 Apply Shape Model to Level Set Method

Now that we have the shape model, we can start to introduce the proposed method. The update equation of the φ function is a little different when we are using shape model which can be written as:

$$\Delta\varphi_{WithShape} = \delta(\varphi)\left(\mu \cdot div(\frac{\nabla\varphi}{|\nabla\varphi|}) - \lambda_1(u_0 - c_1)^2 + \lambda_2(u_0 - c_2)^2 - v + \eta\varphi_S\right) \tag{7}$$

where φ_s is the distance to the contour of shape, and η is the weight of the shape term. Finally, we make a summary of the proposed method.

4 Experimental Results

In order to prove that the proposed shape-based level set method can work well on breast mass segmentation of MRI, we use four images from one patient to test our proposed method, as shown in Fig. 6 (a) and Fig. 6 (d). The image size is 512 × 512.

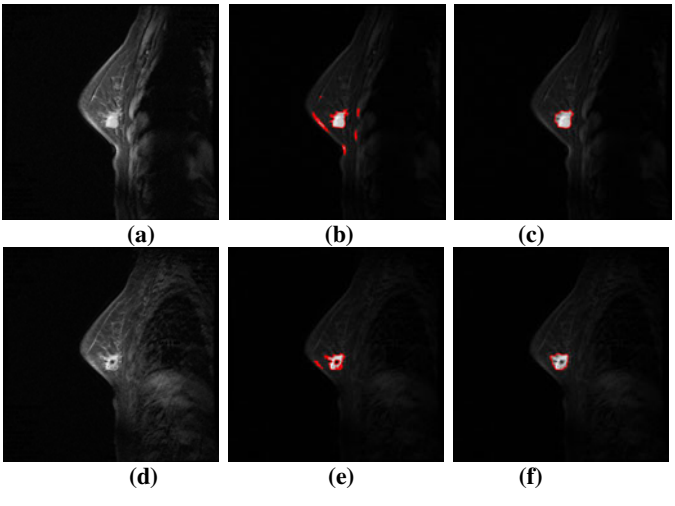


Fig. 6. The segmentation results, (a)(d) Testing breast CT MRIs (b)(e) The segmentation results using Chan-Vese model (c)(f) The segmentation results using the proposed shape-based level set method

The proposed method is compared with the Chan-Vese model. For the level set part, the initialization of φ function is set as circle function where the center of the circle is located on the center of the image and the size of the circle is half of the whole image size, i.e. $\varphi_0(x, y) = -\sqrt{(x-256)^2 + (y-256)^2} + 204.26$. In order to avoid that the φ value diverges, we also normalize the φ value to 1 ~ -1 in the φ function

initialization step. During the φ value updating process, we set $\mu=2\times255^2$, $\lambda_1=1$, $\lambda_2=1$ and $v=-0.4$. After 20 iterations of updates, the level set function will converge. The segmentation results using Chan-Vese model are shown in Fig. 6 (b) and Fig. 6 (e). Figure 6 (c) and Fig. 6 (e) show segmentation results using the proposed method where we set $\eta=1$.

5 Conclusion

In this paper, we proposed a segmentation algorithm to segment the breast mass on the breast MRIs by using level set method incorporated with shape model. In the experiment, we use two breast MRIs to test our segmentation algorithm, and the experimental results show that the proposed shape-based level set method works better on breast mass segmentation comparing with other models. In the future, our proposed method can be applied to many other image segmentation domains, like medical image, temperature image, etc.

Acknowledgement

This work was supported by the National Science Council under grant NSC 99-2221-E-309-008.

References

1. Rao, M., Stough, J., Chi, Y.Y., Muller, K., Tracton, G., Pizer, S.M., Chaney, E.L.: Comparison of human and automatic segmentations of kidneys from CT images. Intl. Journal of Radiation Oncology 61(3), 954–960 (2005)
2. Gunn, S.: On the discrete representation of the Laplacian of Gaussian. Pattern Recognition 32(8), 1463–1472 (1999)
3. Joo, S., et al.: Computer-aided diagnosis of solid breast nodules: use of an srtificial neural network based on multiple sonographic features. IEEE TMI 23(10), 1292–1300 (2004)
4. Manousakas, I.N., Undrill, P.E., Cameron, G.G., Redpath, T.W.: Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions. Computer Biomed. Res. 31, 393–412 (1998)
5. Grau, V., Mewes, A.U.J., Alcaniz, M., Kikinis, R., Warfield, S.K.: Improved watershed transform for medical image segmentation using prior information. IEEE Trans. Medical Imaging 23(4), 447–458 (2004)
6. Chan, T., Vese, L.: Active contours without edges. IEEE Trans. Image Processing 10(2), 266–277 (2001)
7. Cremers, D.: Nonlinear Dynamical Shape Priors for Level Set Segmentation. Journal of Scientific Computing 35(2-3), 132–143 (2008)
8. Chan, T., Vese, L.: A multiphase level set framework for image segmentation using the mumford and shah model. Intl. Journal of Computer Vision 50(3), 271–293 (2002)
9. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Intl. Journal of Computer Vision 1, 321–331 (1988)
10. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Intl. Journal of Computer Vision 22(1), 61–79 (1997)

A New Fast Region Filling Algorithm Based on Cross Searching Method

Feng Duo-le and Zhao Ming

College of Information and Electrical Engineering China Agriculture University
Beijing, China, 100083
fengduole@163.com

Abstract. Region filling algorithm is widely used in GIS, remote sensing, image processing, and plays an important role. There are problems that seeds are filled into stacks repeatedly and redundant information is for traditional 4-connected region filling algorithm. In this paper, we present a new fast region filling algorithm based on "Marking Method" to solve these problems that seeds are filled into stacks repeatedly, and then, change the scanning order and uses the "Cross Method" to search the filling Region. Compared with the traditional algorithm, this algorithm not only can improve 7 times more efficiency, but also can make it simple and filling correctly.

Keywords: computer graphics; region filling; seeds; stact.

1 Introduction

Region-Filling is a method to fill a bounded region with a certain color or image. It means filling the color in a pixel which is connected with other 4 or 8 pixels, and then algorithm fills the color in all pixels in the Region. It is a very important basic algorithm of Computer Graphics as for that it has a wide range of applications in Computer Aided Design, Realistic graphics, Geographic Information System, Image processing, and so on. The adaptability and efficiency of the filling algorithm are highly significant for the practical applications.

Nowadays, the commonly used algorithms are these: scan-line algorithm, edge table of polygon algorithm, seed-filling algorithm as well as polygon edge flog algorithm. These algorithms have some obvious advantages and disadvantages, thus, they do have certain limitations in practical applications. For example, if we do not consider the relevance of pixels, every pixel may enter the stack repeatedly, which will lead to the reducing of efficiency. Besides, this algorithm requires lager stack, so that the stack may be overflow when we fill a large region in this way.

Faced with the drawbacks of the traditional filling algorithm, many scholars have proposed various algorithm improvements, such as reducing stack operations. And some foreign scholars have proposed a kind of seed-filling algorithm based on the chain codes. However, such algorithm not only needs to judge the direction of chain codes, but also has to use Backtracking Algorithm, rather than one-time search. Therefore, this algorithm is too complicated. Based on the previous studies and on the

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 380–387, 2011.

premise of ensuring the correctness of algorithm, this paper will introduce a way to further develop the efficiency of time and space, via different searching solutions to reduce the number of times that the pixels enter the stacks.

1.1 Ideas of Traditional Algorithm

Let's reviews traditional seed filling algorithm: assuming that the algorithm has known the topologies of all of the points and edges of that given polygon. Besides, a certain point in the polygon is knows too, which is called the seed point. Via the seed point, the algorithm begins searching all the inside region of the polygon and fills the color.

The algorithm is divided into 3 steps (Defaulting the algorithm is applied in 4-connected Region):

1. The seed point is pushed into the stack;
2. Performing the following 3-step operation:
   a) Push the top point in the stack out;
   b) Fill color in this point;
   c) Check the 4 points which are connected with this point, if one of the 4 points isn`t on the edge and not filled, push the point in the stack;
3. Check whether the stack is empty. If it is empty, do the 2nd step repeatedly, otherwise stop the algorithm.

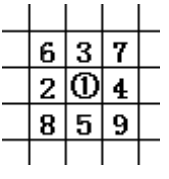


Fig. 1. Schematic diagram of the traditional seed filling algorithm

1.2 Analyzing of the Defect of the Traditional Algorithm

Figure 1 shows the shortages of the traditional filling algorithm.

Assume that point ① is the initial seed point. At this time, the stack is empty. When algorithm starts, point ① is pushed in the stack. And then, point ① is pushed out of the stack and filled the color. At the same time, point ② ③, ④, ⑤ are pushed in. when point ②, ③, ④, ⑤ are out, point ① will enter the stack 4 times repeatedly. And when point ② is out, point ⑥, ⑧ will be pushed in; when point ③ is out, point ⑦, ⑨ will be pushed in. It can be seen from here: point ⑥, ⑦, ⑧, ⑨ will also be pushed in the stack repeatedly.

It is not difficult to find out: all the stack operations after the 1st time are useless. It caused a great waste of system resource that every point will be pushed in the stack 4 times at most. If the algorithm could reduce these excess operations, the efficiency of it will ascend a grade.

2 Improvements of Algorithm

2.1 Ideas of the Improvements

The core of the seed-filling algorithm is depth-first searching. In a certain period of time, thronging all the points in the polygon at least once will complete filling. However, traditional algorithm not only goes through each point once, but also performs stack operations repeatedly. To remedy defects, a "pruning" strategy is presented here – "Marking Method". "Marking Method" will mark another color in points which are already in the stack. Before each point being pushed in the stack, "Marking Method" will check the point whether is marked. If it was marked, algorithm will skip next operations. If not, it will be pushed in the stack and be marked. In this way, "Marking Method" can ensure that each point has one and only one stack operation.

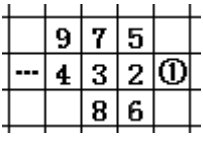


Fig. 2. Schematic diagram of the improved algorithm

2.2 The Steps of Improvements

Figure 2 is a Schematic diagram of the new algorithm.

In order to simplify the question, we only discuss the left side of point ①, because the "Marking Method" can solve the problem of repeated stack operations.

Point ① is pushed out and be filled the color, then point ② is pushed in. After point ② is pushed out, point ⑤, ⑥, ③ enter the stack. After point ③ is pushed out, point ⑦, ⑧, ④ are pushed in……Then the algorithm won't stop and will be performed correctly, until the polygon is filled.

Observing the order of stack operations carefully, we can easily find out that the algorithm will push the points out of the stack and fill in the color along the left side of point ①, in the order that first point ② and then ③, ④ until it arrives at the edge of the polygon. This feature leads to the core of this paper – "Cross Filling". The new algorithm could improve the efficiency greatly.

3 Analyzing of "Cross Filling" Algorithm

3.1 The Principle of "Cross Filling"

"Cross Filling" is still based on the seed-filling algorithm. The initial condition is only the seed point.

There are 3 steps of "Cross Filling":

1. Pushing out. Push the top point out and fill the color. Take the point as the origin, searching in the order that first left and then up, right and down(Searching left

means searching in the order of point ②, ③, ④, ⑤, ⑥. Searching upwards means searching in the order of point ⑦, ⑧, ⑨. And searching right or down is on the analogy of these).

2. Searching. When "Cross Filling" searches, it judges firstly: whether the point is not in the stack or on the edge. If the two conditions are both true, "Cross Filling" will push the points which are near this point in the stack, and then search the next point in the same side repeatedly. If the two conditions are not both true, searching in this side will be stopped and "Cross Filling" will search the next side.
3. Judging. If the stack is not empty, perform the 2 steps above, otherwise "Cross Filling" is stopped.

3.2 Example of "Cross Filling"

In order to express the principle of "Cross Filling" clearly, here is an example of "Cross Filling" in Figure 3.

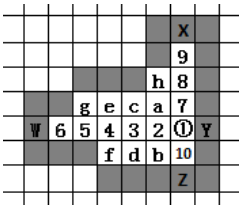


Fig. 3. The principle of algorithm

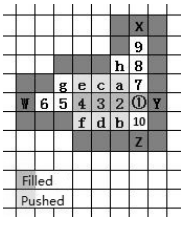


Fig. 4. Process of "Cross Filling"1

Point ① is pushed out and be filled;

Take point ① as the origin and search from the origin;

Search the left side of point ①;

Visit point ②, ③, ④ and fill them color in order, and Point a ~ Point f are pushed in the stack successively (Figure 4) ;

Visit point ⑤ and fill it color, and only one point (Point g) is available, so push it in the stack;

Visit point ⑥ and fill it color, but there are no points available;

Visit Point W and W is on the edge. Therefore, stop searching in the left side;

Search the upwards of point ①;

Visit point ⑦, ⑧, ⑨ and fill them color in order. After that Point h is pushed in the stack (Point a has already been in the stack);

Visit Point X and X is on the edge, so stop search in this direction;

Search the point above point ①;

Visit Point Y and Y is on the edge, so stop search in right direction;

Search the down of point ①;

Visit point ⑩ and fill it color, but there are no points available;

Visit Point Z and Z is on the edge, stop search in down direction (Figure 5);

Searching finished from point ①. Now, Point h is pushed out and is filled.

Take Point h as the origin and search from the origin: there is no available point next to Point h.

Push out all the points in order and fill them color, and then "Cross Filling" has finished.

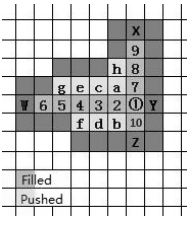


Fig. 5. Process of "Cross Filling"2

3.3 Theoretical Analysis of "Cross Filling"

The core of "Cross Filling" is deep first searching. It fills all the region in such circumstances that not finding for edges or the intersections of edges.

According to the traditional algorithm, each point will be pushed in the stack at least once. Algorithm would rather fill color in points than pushed them in and out. So "Cross Filling" sets the top point in the stack as the origin, in the order that first up and then down, left and right. The advantages of the improvements are reducing a large quantity of points which were pushed in the stack. In the traditional algorithm, when one point is pushed out, algorithm will fill one point. But in "Cross Filling", when one point is pushed out, algorithm will fill a group of points. So "Cross Filling" is faster than the traditional algorithm. In order to reducing the Computational complexity further, "Cross Filling" adds a judge when points are pushed out: if this point has already filled, continue with the next step. We will prove the correctness of this "pruning". We know if a point will be filled, there are only two cases: the one is that this point will be filled when it is pushed out; the other is that there must be points out which are in the same line or row with this point. If the 1st circumstance happens, this point must be not filled before it is out. Therefore, algorithm fills the color in this point and continues to search for the next. The 2nd circumstance will happen, when one of the points which are in the up, down, left or right side of this point is operated to be pushed out of the stack (Figure 6). Assuming that point ① should have been pushed out of the stack, but point ① has been filled, so skip to the next step. But points in the stack which are from point ① did not reduce: point ① and ⑩ are in the same line, so points between point ⑧ and ⑨ will be pushed in the stack, they did not change. Points between d ~ g are pushed in the stack instead of the points between point ⑦ ~ ?. Above this, this "pruning" method is correct.

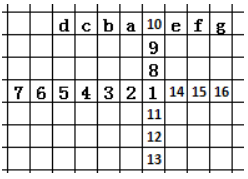


Fig. 6. The proof of the Correctness

4 Results Comparison of Algorithms

In order to verify the correctness and efficient of new algorithms, we did a comprehensive comparison test with traditional algorithm.

4.1 Preparation of the Tests

Involved in the testing algorithms:

The traditional seed-filling algorithm; "Cross Filling".

Table 1 is the environment of tests

Table 1. Test environment

Test environment	
CPU	Intel Core2 Duo P8700 2.53GHz
RAM	4GB DDR3 1066MHz
Operation System	Windows7 64bit Ultimate + SP1
Programming Tool	VisualStudio2010 MFC

Testing requirements:

There are 3 groups need to be tested. The filling region of the first group is a circle whose radius is 100 pixels. The filling region of the second group is a square whose edge length is 100 pixels. The filling region of the third group is a polygon (Figure 7).

For each group, each algorithm will be tested 3 times. The test result is the average of 3 times. The Time-complexity divided into two parts: one is the time of algorithms run (in milliseconds); the other is program step. The Space-complexity is the maximum depth in the algorithm running (if an algorithm did not use stack, did not measure the Space-complexity).

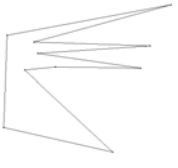


Fig. 7. The polygon of the 3rd group

4.2 The Data Analysis

The following is the format of the test results:

"Time (milliseconds) / Stack Depth (layer) / Program Steps (step)".

For example: "686/438/2304" means that in this group the algorithm cost 686ms, used 438 layers stack and run in 2304 steps.

The table 2 is the test result of group 1 – circle; the table 3 is the test result of group 2 – square; the table 4 is the test result of group 3 – polygon; table 5 is the contrast of the efficiency between the traditional seed-filling algorithm and the "Cross Filling" algorithm.

Table 2. Group 1 – Circle

Algorithms / Times	The traditional seed-filling algorithm	"Cross Filling" algorithm
1st	68656/30306/61013	10421/13293/30703
2nd	68297/30327/61013	10296/13833/30703
3rd	68648/30310/61013	9859/4532/30703
Average	68534/30314/61013	10192/10553/30703

Table 3. Group 2 – Square

Algorithms / Times	The traditional seed-filling algorithm	"Cross Filling" algorithm
1st	90714/38810/78013	12449/19308/39204
2nd	89451/38810/78013	11232/19307/39204
3rd	90403/38810/78013	11902/5439/39204
Average	90189/38810/78013	11861/14685/39204

Table 4. Group 3 – Polygon

Algorithms / Times	The traditional seed-filling algorithm	"Cross Filling" algorithm
1st	110278/47352/96303	17176/9120/48952
2nd	110293/26200/96303	14820/8544/48952
3rd	110012/47328/96303	15896/4575/48952
Average	110194/40293/96303	15964/7413/48952

Table 5. Improvements efficiency of algorithms

	Time of the traditional seed-filling algorithm	Time of "Cross Filling" algorithm	Improvement (%)
Group 1	68534	10192	672.43%
Group 2	90189	11861	760.38%
Group 3	110194	15964	690.27%
Average	268917	38017	707.69%

The results of table 5 showed that the "Cross Filling" algorithm is 7 times faster than the traditional algorithm. Especially, in the regular polygon, the efficiency of the algorithm can achieve the ultimate.

5 Conclusion

At first, "Cross Filling" uses "Marking Method" to solve the problem about repeated stack operations fundamentally. Then the "Cross Searching" is the main reason to improve efficiency. It reduces the number of points onto the stack, shortens the execution time, and reduces the memory usage. Simultaneously, using array as the stack reduces the dynamic allocation of memory space each time wasted.

This fast region filling algorithm is implemented with C++ function module. It is much more practical for the reason that it is more efficient, widely adopted as well as easily to realize or protect. Besides, it has many other advantages, such as filling the color accurately and so on.

Acknowledgements

This research is supported by the grant from curriculum reform Grants of China Agriculture University. The authors are grateful for the anonymous reviewers who made constructive comment.

References

1. Liu, X.B., Hu, F.S., Zhang, B.J.: A new regional seed-filling algorithm. Computer Engineering and Applications 38(8), 101–103 (2002) (in Chinese)
2. Burger, P.: Interactive computer graphices functional, procedural and device-level methods. Addison-Wesley, Massachusetts (1989)
3. Pavlidis, T.: Algorithms for graphics and image processing, pp. 65–73. Computer Science Press Inc., Rockville (1982)
4. Sun, J.G., Hu, S.M.: Computer Graphics Essentials, pp. 24–31. Tsinghua University Press, Beijing (2008) (in Chinese)
5. Sun, Q., Zheng, Y.G.: Scan line fill algorithm's problems and improvements in complicated connected region. Image Processing and Mode Identification (3), 71–75 (2007) (in Chinese)
6. Guo, W.P., Long, B.Q.: The improvements of Scan line seed-filling. Newspapers of Tianjin Polytechnic University 27(2), 48–51 (2008) (in Chinese)
7. Zhang, Z.F., Ma, S.F., Li, W.: A new seed-filling algorithm. Computer Engineering and Applications 45(6), 201–206 (2009) (in Chinese)
8. Dorigo, M., Caro, G.D., Gambardella, L.M.: Ant algorithms for solving weapon-target at assignment problem. Applied Soft. Computing 2, 39–47 (2002)

Energy Demand Forecasting in China Based on Dynamic RBF Neural Network*

Dongqing Zhang, Kaiping Ma, and Yuexia Zhao

College of Engineering, Nanjing Agricultural University, Nanjing, 210031, China
hollycase@163.com, kaiping_ma@126.com, zhaoyuexia@njau.edu.cn

Abstract. A dynamic radial basis function (RBF) network model is proposed for energy demand forecasting in this paper. Firstly, we present a time series forecasting framework based on variable structure RBF network. In this framework, both the number of basis function and the input orders are variable. Secondly, an on-line prediction algorithm using sequential Monte Carlo (SMC) method is developed. Due to the high dimensional state-spaces, the Rao-Blackwellised particle filter is adopted to compute the posterior probability density function of state variables. In this SMC algorithm, the sub-space sampling, state prediction, weight updating, exact computation with Kalman filter and the change of RBF structure have been discussed in detail. At last, the data of total energy demand in China are analyzed and experimental results indicate that the proposed model and prediction algorithm are effective.

1 Introduction

Energy is an important material resource of existence, economy development and modern civilization. It is an essential strategic supplies relation to national economy and defense safety. Therefore, scientific forecast of energy demand has an essential meaning.

Studies on energy demand forecasting of China have began in the 1980s. Traditional forecast methods, such as Scenario analysis, input-output model, co-integration theory, regression analysis, combining forecasts model and so on, are used to predict the energy demand [1]. So far, modern methods, such as neural network, are recommended for the energy demand forecast. Wang [2] proposed the static neural network to predict energy demand in China. In this paper, a dynamic RBF neural network is presented for energy demand forecast and corresponding online prediction is developed. To the best of my knowledge, there are no literatures to discuss the energy demand forecast using dynamic neural network.

The paper is organized as follows. In Section 2, a variable structure RBF neural network for energy demand forecasts is proposed. In Section 3, online prediction based on the variable structure RBF networks using SMC method is studied thoroughly and corresponding algorithm is developed. Section 4 presents the empirical analysis. Finally, a summary is given in Section 5.

* This work is supported by Education Department Foundation of Jiangsu Province (2010SJD630061), Bureau of Agricultural Machinery Foundation of Jiangsu Province (gxz09003), and Science-Technology Foundation for Young Scientists of NAU(KJ09031).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 388–395, 2011.

2 Variable Structure RBF Network Models

Energy demand is affected by many factors, such as GDP, population, import, export, employment, and so on, therefore it shows obviously nonlinearity, refer to Fig.1. The data come from National Bureau of Statistics of China (http://www.stats.gov.cn/). Suppose $y_1,\cdots,y_{t-1}$ are the actual energy demand in history, the predictive value $\hat{y}_t$ at time *t* can be obtained

$$\hat{y}_t = f(y_{t-1},\cdots,y_{t-g_t},e_{t-1},\ldots,e_{t-q_t}) + e_t \tag{1}$$

Where g_t, q_t are the order of the autoregressive part and moving part respectively, which vary along with time. $\{e_{t-i}, i=0,\cdots,q_t\}$ is measurement noise, which is generally unobserved and can be replaced by prediction error, i.e. $e_{t-i} = y_{t-i} - \hat{y}_{t-i}, i=1,\cdots,q_t$.

According to (1), the key for prediction is to choose proper expression of nonlinear function $f(\cdot)$. However, in many cases, it is impossible to express nonlinear function in exact form, so many researchers resort to neural networks, such as feedforward neural network [3] and dynamic RBF models with the variable number of basis functions [4] and so on. In this paper, a variable structure RBF networks model, in which both the number of radial basis functions and input order change along with time, is chosen to approximate nonlinear function.

We shall consider an approximation scheme consisting of a mixture of k_t radial basis functions and a linear regression term proposed by Holmes in 1998. More precisely, the linear-RBF model M at time t is given by [5]

$$\hat{y}_t = \sum_{j=1}^{k_t} a_{j,t}\phi(\| \boldsymbol{x}_t - \boldsymbol{c}_{j,t} \|) + b_t + \boldsymbol{\beta}_t^T \boldsymbol{x}_t + e_t,\ k_t \geq 0 \tag{2}$$

where $\boldsymbol{x_t} = (y_{t-g_t},\cdots,y_{t-1},e_{t-q_t},\cdots,e_{t-1})^T$ is the input of RBF network, $\|.\|$ denotes a distance metric, $\phi(x) = e^{-\lambda x^2}$ is the basis functions, $\boldsymbol{c}_{j,t}$ denotes the *j* th RBF centre with k_t RBFs, $a_{j,t}$ denotes the *j* th RBF amplitude, b_t and $\boldsymbol{\beta}_t = (\beta_{1,t},\cdots,\beta_{d_t,t})^T$ are the linear regression parameters, and $d_t = g_t + q_t$. For convenience, (2) is expressed as

$$\hat{y}_t = \boldsymbol{D}_t\boldsymbol{\alpha}_t + e_t \tag{3}$$

where $\boldsymbol{D_t} = [1, y_{t-g_t},\cdots, y_{t-1}, e_{t-q_t},\cdots,e_{t-1}, \phi(\boldsymbol{x_t},\boldsymbol{c}_{1,t}),\cdots,\phi(\boldsymbol{x_t},\boldsymbol{c}_{k_t,t})]$, $\boldsymbol{\alpha_t} = [b_t,\beta_{1,t},\cdots,\beta_{d_t,t},a_{1,t},\cdots,a_{k_t,t}]^T$, the noise process is assumed to be normally distributed as follows: $e_t \sim N(0,\sigma_t^2)$.

3 On-Line Prediction Using SMC Methods

In this section, our intention is to obtain $\hat{y}_t$ using variable structure RBF networks model given all available information $y_{1:t-1} = (y_1,\cdots,y_{t-1})$.

3.1 State Transition Equation

The state transition equations of the proposed model are given by

$$\begin{cases} k_t \sim P(k_t \mid k_{t-1}),\ g_t \sim P(g_t \mid g_{t-1}),\ q_t \sim P(q_t \mid q_{t-1}) \\ \boldsymbol{c}_t = \boldsymbol{c}_{t-1} + \boldsymbol{\varepsilon}_{\boldsymbol{c}},\ \boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \boldsymbol{\varepsilon}_{\boldsymbol{\alpha}},\ \ln(\sigma_t^2) = \ln(\sigma_{t-1}^2) + \varepsilon_\sigma \end{cases} \tag{4}$$

where $P(k_t \mid k_{t-1}), P(g_t \mid g_{t-1}), P(q_t \mid q_{t-1})$ are specified discrete distributions, $\boldsymbol{\varepsilon}_{\boldsymbol{c}} \sim N(\boldsymbol{0}, \delta_{\boldsymbol{c}}^2 \boldsymbol{I}_{d_t \times d_t})$, $\varepsilon_\sigma \sim N(0, \delta_\sigma^2), \boldsymbol{\varepsilon}_{\boldsymbol{\alpha}} \sim N(\boldsymbol{0}, \delta_{\boldsymbol{\alpha}}^2 \boldsymbol{I}_{m_t \times m_t})$, and $m_t = k_t + d_t + 1$. $\delta_{\boldsymbol{c}}^2, \delta_{\boldsymbol{\alpha}}^2, \delta_\sigma^2$ are constants.

3.2 Rao-Blackwellised Particle Filter

Consider the variable structure RBF networks model, the posterior pdf $p(\boldsymbol{\theta}_t \mid y_{1:t})$ should be calculated recursively, where $\boldsymbol{\theta}_t = \{k_t, \boldsymbol{\alpha}_t, \boldsymbol{c}_t, g_t, q_t, \sigma_t^2\}$. If the posterior pdf is a linear Gaussian distribution, it can be solved with analytical methods such as Kalman filter. However, on most occasions, it is impossible to derive closed-form analytical expressions to obtain the posterior distribution and its features of interest. To compute these distributions in a reasonable time, we need to resort to SMC method. But SMC method in high dimensional state-spaces can be inefficient because a large number of samples are needed to represent the posterior. Therefore, we adopt Rao-Blackwellised particle filter (RBPF) [6], which basic idea is to partition the state-space into a few sub-spaces and to compute some of the variables exactly and sample the rest.

Suppose we partition the state-space into two sub-spaces, i.e. $\boldsymbol{\theta}_t = \{\boldsymbol{r}_t, \boldsymbol{s}_t\}$, where $\boldsymbol{r}_t = \{k_t, \boldsymbol{c}_t, g_t, q_t, \sigma_t^2\}$, $\boldsymbol{s}_t = \{\boldsymbol{\alpha}_t\}$. Then, we can write

$$p(\boldsymbol{r}_t, \boldsymbol{s}_t \mid y_{1:t}) = p(\boldsymbol{r}_t \mid y_{1:t}) p(\boldsymbol{s}_t \mid \boldsymbol{r}_t, y_{1:t}) \tag{5}$$

We sample $p(\boldsymbol{r}_t \mid y_{1:t})$ using standard SMC method. It is clear that, given $\boldsymbol{r}_t$, we end up with a linear Gaussian system with states $\boldsymbol{\alpha}_t$, i.e. $p(\boldsymbol{\alpha}_t \mid \boldsymbol{r}_t, y_{1:t}) = N(\overline{\boldsymbol{\alpha}}_t, \boldsymbol{P}_t)$, where $\overline{\boldsymbol{\alpha}}_t = E\{\boldsymbol{\alpha}_t\}$ and $\boldsymbol{P}_t = \mathrm{cov}\{\boldsymbol{\alpha}_t, \boldsymbol{\alpha}_t\} = E\{(\boldsymbol{\alpha}_t - \overline{\boldsymbol{\alpha}}_t)(\boldsymbol{\alpha}_t - \overline{\boldsymbol{\alpha}}_t)^T\}$ correspond to the conditional mean and covariance matrix of $\boldsymbol{\alpha}_t$ respectively, so we can update $p(\boldsymbol{s}_t \mid \boldsymbol{r}_t, y_{1:t})$ analytically and efficiently using Kalman filter.

3.3 On-Line Prediction Algorithm

Suppose that we have

$$p(\boldsymbol{r}_{t-1} \mid y_{1:t-1}) \approx \sum_{i=1}^{n_p} w_{t-1}^{(i)} \delta(\boldsymbol{r}_{t-1} - \boldsymbol{r}_{t-1}^{(i)}) \tag{6}$$

$$p(\boldsymbol{s}_{t-1} \mid \boldsymbol{r}_{t-1}^{(i)}, y_{1:t-1}) = N(\overline{\boldsymbol{\alpha}}_{t-1}^{(i)}, \boldsymbol{P}_{t-1}^{(i)}),\quad i = 1, \cdots n_p \tag{7}$$

where $\{\boldsymbol{r}_{t-1}^{(i)}\}_{i=1}^{n_p}$ is a set of particles with weights $\{w_{t-1}^{(i)}\}_{i=1}^{n_p}$, n_p is the number of particles, $\delta(\cdot)$ is delta function. Now we will calculate $p(\hat{y}_t \mid y_{1:t-1})$ and $p(\boldsymbol{r}_t, \boldsymbol{s}_t \mid y_{1:t})$.

3.3.1 Sampling Step

Because

$$p(\boldsymbol{r}_t \mid y_{1:t-1}) = \int p(\boldsymbol{r}_t, \boldsymbol{r}_{t-1} \mid y_{1:t-1}) d\boldsymbol{r}_{t-1} = \int p(\boldsymbol{r}_{t-1} \mid y_{1:t-1}) p(\boldsymbol{r}_t \mid r_{t-1}, y_{1:t-1}) d\boldsymbol{r}_{t-1} \tag{8}$$

Substituting (6) into (8) yields

$$p(\boldsymbol{r}_t \mid y_{1:t-1}) \approx \int \sum_{i=1}^{n_p} w_{t-1}^{(i)} \delta(\boldsymbol{r}_t - \boldsymbol{r}_t^{(i)}) p(\boldsymbol{r}_t \mid r_{t-1}, y_{1:t-1}) d\boldsymbol{r}_{t-1} = \sum_{i=1}^{n_p} w_{t-1}^{(i)} p(\boldsymbol{r}_t \mid \boldsymbol{r}_{t-1}^{(i)}, y_{t-1}) \tag{9}$$

We adopt $p(\boldsymbol{r}_t \mid \boldsymbol{r}_{t-1}^{(i)}, y_{t-1}) = p(\boldsymbol{r}_t \mid \boldsymbol{r}_{t-1}^{(i)})$ as the important density function, namely, draw new particles $\boldsymbol{r}_t^{(i)} = \{k_t^{(i)}, \boldsymbol{c}_t^{(i)}, g_t^{(i)}, q_t^{(i)}, \sigma_t^{2(i)}\}$ according to (4). Then we have

$$p(\boldsymbol{r}_t \mid y_{1:t-1}) \approx \sum_{i=1}^{n_p} w_{t-1}^{(i)} \delta(\boldsymbol{r}_t - \boldsymbol{r}_t^{(i)}) \tag{10}$$

3.3.2 Measurement Prediction Step

By the Chapman-Kolmogorov equation and (10), we have

$$p(\hat{y}_t \mid y_{1:t-1}) = \int p(\boldsymbol{r}_t \mid y_{1:t-1}) p(\hat{y}_t \mid \boldsymbol{r}_t, y_{1:t-1}) d\boldsymbol{r}_t = \sum_{i=1}^{n_p} w_{t-1}^{(i)} p(\hat{y}_t \mid \boldsymbol{r}_t^{(i)}, y_{1:t-1}) \tag{11}$$

Obviously, once $\boldsymbol{r}_t^{(i)}$ has been drawn, $\boldsymbol{D}_t^{(i)}$ can be obtained according to its definition. Because $\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \boldsymbol{\varepsilon}_{\boldsymbol{\alpha}}$, $\hat{y}_t = \boldsymbol{D}_t \boldsymbol{\alpha}_t + e_t$, and $e_t \sim N(0, \sigma_t^2)$, thus

$$p(\boldsymbol{\alpha}_t \mid \boldsymbol{r}_t^{(i)}, y_{1:t-1}) = N(\overline{\boldsymbol{\alpha}}_{t-1}^{(i)}, \boldsymbol{P}_{t-1}^{(i)} + \delta_{\boldsymbol{\alpha}}^2 \boldsymbol{I}_{m_t \times m_t}),\ i = 1, \cdots, n_p \tag{12}$$

$$p(\hat{y}_t \mid \boldsymbol{r}_t^{(i)}, y_{1:t-1}) = N(\boldsymbol{D}_t^{(i)} \overline{\boldsymbol{\alpha}}_{t-1}^{(i)}), \boldsymbol{D}_t^{(i)} (\boldsymbol{P}_{t-1}^{(i)} + \delta_{\boldsymbol{\alpha}}^2 \boldsymbol{I}_{m_t \times m_t}) (\boldsymbol{D}_t^{(i)})^T + \sigma_t^{2(i)})),\ i = 1, \cdots, n_p \tag{13}$$

Substituting (13) into (11) then yields

$$p(\hat{y}_t \mid y_{1:t-1}) \approx \sum_{i=1}^{n_p} w_{t-1}^{(i)} N(\boldsymbol{D}_t^{(i)} \overline{\boldsymbol{\alpha}}_{t-1}^{(i)}, \boldsymbol{D}_t^{(i)} (\boldsymbol{P}_{t-1}^{(i)} + \delta_{\boldsymbol{\alpha}}^2 \boldsymbol{I}_{m_t \times m_t}) (\boldsymbol{D}_t^{(i)})^T + \sigma_t^{2(i)}) \tag{14}$$

In real application, the mean of $p(\hat{y}_t \mid y_{1:t-1})$ is often used to denote the measurement prediction, that is $\hat{y}_t = \sum_{i=1}^{n_p} w_{t-1}^{(i)} \boldsymbol{D}_t^{(i)} \overline{\boldsymbol{\alpha}}_{t-1}^{(i)}$.

3.3.3 Weight Updating and Resampling Step

When the true measurement y_t is available, we can update the weight according to $w_t^{(i)} = w_{t-1}^{(i)} p(y_t \mid \boldsymbol{r}_t^{(i)}, y_{1:t-1})$ [7], where the likelihood $p(y_t \mid \boldsymbol{r}_t^{(i)}, y_{1:t-1})$ has been presented in (13). Obviously, all particles have different weights. If some particles have too small weights, it will cause degeneracy. In order to avoid the degeneracy

phenomenon, systematic resampling is preferred by author. For more details about systematic resampling, refer to [8]. Thus we have

$$p(\boldsymbol{r}_t \mid y_{1:t}) \approx \sum_{i=1}^{n_p} w_t^{(i)} \delta(\boldsymbol{r}_t - \boldsymbol{r}_t^{(i)}) \tag{15}$$

3.3.4 Exact Step

Now we will study $p(\boldsymbol{s}_t \mid \boldsymbol{r}_t, y_{1:t})$. Because $\boldsymbol{s}_t$ is conditionally linear Gaussian given $\boldsymbol{r}_t$, it is possible to integrate out $\boldsymbol{s}_t$, thereby reducing the variance of the estimates. The end result corresponds to a bank of Kalman filters for network output:

$$\begin{cases} \overline{\boldsymbol{\alpha}}_t^{(i)} = \overline{\boldsymbol{\alpha}}_{t-1}^{(i)} + \boldsymbol{K}_t^{(i)}(y_t - \boldsymbol{D}_t^{(i)}\overline{\boldsymbol{\alpha}}_{t-1}^{(i)}) \\ \boldsymbol{P}_t^{(i)} = \boldsymbol{P}_{t-1}^{(i)} + \delta_{\alpha}^2 \boldsymbol{I}_{m_t \times m_t} - \boldsymbol{K}_t^{(i)} \boldsymbol{D}_t^{(i)} (\boldsymbol{P}_{t-1}^{(i)} + \delta_{\alpha}^2 \boldsymbol{I}_{m_t \times m_t}) \end{cases} \tag{16}$$

where $\boldsymbol{K}_t^{(i)} = (\boldsymbol{P}_{t-1}^{(i)} + \delta_{\alpha}^2 \boldsymbol{I}_{m_t \times m_t})(\boldsymbol{D}_t^{(i)})^T [\sigma_t^{2(i)} + \boldsymbol{D}_t^{(i)} (\boldsymbol{P}_{t-1}^{(i)} + \delta_{\alpha}^2 \boldsymbol{I}_{m_t \times m_t})(\boldsymbol{D}_t^{(i)})^T]^{-1}$ is Kalman gain.

According to (17), we obtain $\overline{\boldsymbol{\alpha}}_t^{(i)}$ and $\boldsymbol{P}_t^{(i)}$, such that

$$p(\boldsymbol{s}_t \mid \boldsymbol{r}_t^{(i)}, y_{1:t}) = p(\boldsymbol{\alpha}_t \mid \boldsymbol{r}_t^{(i)}, y_{1:t}) = N(\overline{\boldsymbol{\alpha}}_t^{(i)}, \boldsymbol{P}_t^{(i)}), \quad i = 1, \cdots, n_p \tag{17}$$

From above discussion, we obtain $\hat{y}_t$ and $p(\boldsymbol{r}_t, \boldsymbol{s}_t \mid y_{1:t}) = p(\boldsymbol{r}_t \mid y_{1:t}) p(\boldsymbol{s}_t \mid \boldsymbol{r}_t, y_{1:t})$ recursively.

3.4 Parameters Adjustment

In this section, some parameters will be adjusted along with k_t, g_t, q_t. For convenience, $\boldsymbol{D}_t, \boldsymbol{\alpha}_t, \boldsymbol{P}_t$ are represented as

$$\begin{aligned} &\boldsymbol{D}_t = (\boldsymbol{D}_{t,1}, \boldsymbol{D}_{t,2}, \boldsymbol{D}_{t,3}),\ \boldsymbol{D}_{t,1} = 1,\ \boldsymbol{D}_{t,2} = \boldsymbol{x}_t^T,\ \boldsymbol{D}_{t,3} = (\phi(\boldsymbol{x}_t, \boldsymbol{c}_{1,t}), \cdots, \phi(\boldsymbol{x}_t, \boldsymbol{c}_{k_t,t})) \\ &\boldsymbol{\alpha}_t = (\boldsymbol{\alpha}_{t,1}, \boldsymbol{\alpha}_{t,2}, \boldsymbol{\alpha}_{t,3})^T,\ \boldsymbol{\alpha}_{t,1} = b_t,\ \boldsymbol{\alpha}_{t,2} = (\beta_{1,t}, \cdots, \beta_{d_t,t}),\ \ \boldsymbol{\alpha}_{t,3} = (a_{1,t}, \cdots, a_{k_t,t}) \\ &\boldsymbol{P}_t = (\boldsymbol{P}_{t,i,j})_{i,j=1,2,3},\ \boldsymbol{P}_{t,i,j} = \mathrm{cov}\{\boldsymbol{\alpha}_{t,i}^T, \boldsymbol{\alpha}_{t,j}^T\} = E\{(\boldsymbol{\alpha}_{t,i} - \overline{\boldsymbol{\alpha}}_{t,i})^T (\boldsymbol{\alpha}_{t,j} - \overline{\boldsymbol{\alpha}}_{t,j})\} \end{aligned} \tag{18}$$

For distinguishing, we use the subscripts b, a to represent the parameters adjusted before and after respectively in this section.

3.4.1 Change of the Number of Basis Functions

When k_t changes, the dimensions of $\boldsymbol{D}_t, \boldsymbol{\alpha}_t, \boldsymbol{P}_t$ will change accordingly. For simplicity, we only consider the following three cases: (i) $k_t = k_{t-1}$; (ii) $k_t = k_{t-1} - 1$; (iii) $k_t = k_{t-1} + 1$. As for case (i), we needn't adjust parameters any more and only study the rest two cases. If k_t decreases, one node will be deleted which is uniformly selected from the existing set of basis functions [9]. Subsequently delete the corresponding terms in $\boldsymbol{D}_t, \boldsymbol{\alpha}_t$ and the corresponding row and column entries in the matrix $\boldsymbol{P}_t$. If k_t increases, i.e $k_t = k_{t-1} + 1$, we can modify $\boldsymbol{D}_t, \boldsymbol{\alpha}_t, \boldsymbol{P}_t$ as follows

$$\begin{cases} \boldsymbol{D}_{t,1}^a = \boldsymbol{D}_{t,1}^b,\ \boldsymbol{D}_{t,2}^a = \boldsymbol{D}_{t,2}^b,\ \boldsymbol{D}_{t,3}^a = (\boldsymbol{D}_{t,3}^b, \phi(\boldsymbol{x}_t, \boldsymbol{c}_{\boldsymbol{k_t}+1,t})) \\ \boldsymbol{\alpha}_{t,1}^a = \boldsymbol{\alpha}_{t,1}^b,\ \boldsymbol{\alpha}_{t,2}^a = \boldsymbol{\alpha}_{t,2}^b,\ \boldsymbol{\alpha}_{t,3}^a = (\boldsymbol{\alpha}_{t,3}^b, a_{k_t+1,t}),\ \boldsymbol{P}_t^a = diag\{\boldsymbol{P}_t^b, \eta\} \end{cases} \tag{19}$$

where new basis center $\boldsymbol{c}_{k_t+1,t}$ is sampled from a Gaussian distribution centered at $\boldsymbol{x}_t$, $a_{k_t+1,t}$ and η depend on our prior knowledge about the data set being modeled.

3.4.2 Change of Input Order

The dimensions of $\boldsymbol{x}_t, \boldsymbol{c}_{i,t}, \boldsymbol{D}_t, \boldsymbol{\alpha}_t, \boldsymbol{P}_t$ and some entries in $\boldsymbol{D}_t, \boldsymbol{\alpha}_t, \boldsymbol{P}_t$ will change along with input order g_t and q_t. We first discuss the case of g_t. For simplicity, we only study three cases: (i) $g_t = g_{t-1}$;(ii) $g_t = g_{t-1} + 1$; (iii) $g_t = g_{t-1} - 1$. As for case (i), the parameters are needn't adjusted any more, and only the rest two cases need to be studied. If g_t increases, namely, add y_{t-g_t-1} into the input data, we can extend $\boldsymbol{x}_t, \boldsymbol{c}_t, \boldsymbol{D}_t$ as follows.

$$\begin{cases} \boldsymbol{x}_t^a = (y_{t-g_t-1}, (\boldsymbol{x}_t^b)^T)^T, \quad \boldsymbol{c}_{i,t}^a = (c_{i,t,0}, \boldsymbol{c}_{i,t}^b),\ 1 \le i \le k_t \\ \boldsymbol{D}_{t,1}^a = \boldsymbol{D}_{t,1}^b,\ \boldsymbol{D}_{t,2}^a = (y_{t-g_t-1}, \boldsymbol{D}_{t,2}^b),\ \boldsymbol{D}_{t,3}^a = (\phi(\boldsymbol{x}_t^a, \boldsymbol{c}_{1,t}^a), \cdots, \phi(\boldsymbol{x}_t^a, \boldsymbol{c}_{k_t,t}^a)) \end{cases} \tag{20}$$

where $\{c_{i,t,0}, i = 1, \cdots, k_t\}$ is sampled from a Gaussian distribution centered at y_{t-g_t-1} .

Because $\phi(\boldsymbol{x}_t^a, \boldsymbol{c}_{i,t}^a) = \exp(-\lambda(y_{t-g_t-1} - c_{i,t,0})^2)\phi(\boldsymbol{x}_t^b, \boldsymbol{c}_{i,t}^b), i = 1, \cdots, k_t$, then $\boldsymbol{D}_{t,3}^a = \boldsymbol{D}_{t,3}^b \boldsymbol{U}^{-1}$, where $\boldsymbol{U} = diag\{\exp(\lambda(y_{t-g_t-1} - c_{1,t,0})^2), \cdots, \exp(\lambda(y_{t-g_t-1} - c_{k_t,t,0})^2)\}$. By the consistency of signal projection, there is $\boldsymbol{D}_{t,3}^a(\boldsymbol{\alpha}_{t,3}^a)^T = \boldsymbol{D}_{t,3}^b \boldsymbol{U}^{-1}(\boldsymbol{\alpha}_{t,3}^a)^T = \boldsymbol{D}_{t,3}^b(\boldsymbol{\alpha}_{t,3}^b)^T$. So $\boldsymbol{\alpha}_t^a$ can be modified as

$$\boldsymbol{\alpha}_{t,1}^a = \boldsymbol{\alpha}_{t,1}^b,\ \boldsymbol{\alpha}_{t,2}^a = (\beta_{0,t}, \boldsymbol{\alpha}_{t,2}^b),\ \boldsymbol{\alpha}_{t,3}^a = \boldsymbol{\alpha}_{t,3}^b \boldsymbol{U}^T \tag{21}$$

where $\beta_{0,t}$ is generated in random. Since $\boldsymbol{P}_t = \mathrm{cov}\{\boldsymbol{\alpha}_t, \boldsymbol{\alpha}_t\}$, we have

$$\begin{cases} \boldsymbol{P}_{t,1,1}^a = \boldsymbol{P}_{t,1,1}^b,\ \boldsymbol{P}_{t,1,2}^a = (\boldsymbol{0}, \boldsymbol{P}_{t,1,2}^b),\ \boldsymbol{P}_{t,1,3}^a = \boldsymbol{P}_{t,1,3}^b \boldsymbol{U}^T \\ \boldsymbol{P}_{t,2,1}^a = (\boldsymbol{0}, (\boldsymbol{P}_{t,2,1}^b)^T)^T,\ \boldsymbol{P}_{t,2,2}^a = diag\{\eta, \boldsymbol{P}_{t,2,2}^b\},\ \boldsymbol{P}_{t,2,3}^a = (0, \boldsymbol{U}(\boldsymbol{P}_{t,2,3}^b)^T)^T \\ \boldsymbol{P}_{t,3,1}^a = \boldsymbol{U}\boldsymbol{P}_{t,3,1}^b,\ \boldsymbol{P}_{t,3,2}^a = (0, \boldsymbol{U}\boldsymbol{P}_{t,3,2}^b),\ \boldsymbol{P}_{t,3,3}^a = \boldsymbol{U}\boldsymbol{P}_{t,2,3}^b \boldsymbol{U}^T \end{cases} \tag{22}$$

If g_t decreases, namely, delete the input y_{t-g_t} , so we delete the corresponding terms in $\boldsymbol{x}_t, \boldsymbol{c}_t, \boldsymbol{D}_t, \boldsymbol{\alpha}_t$ and the corresponding row and column entries in the matrix $\boldsymbol{P}_t$, besides, $\boldsymbol{D}_t, \boldsymbol{\alpha}_t, \boldsymbol{P}_t$ should be modified as the case of that g_t increase. Note that the matrix $\boldsymbol{U}$ in $\boldsymbol{D}_t^a, \boldsymbol{\alpha}_t^a, \boldsymbol{P}_t^a$ will be modified as

$$\boldsymbol{U} = diag\{\exp(-\lambda(y_{t-g_t-1} - c_{1,t,0})^2), \cdots, \exp(-\lambda(y_{t-g_t-1} - c_{k_t,t,0})^2)\} \tag{23}$$

As for the change of q_t , the analysis process is similar to that of g_t , so we omit it.

4 Experiment Researches

In this section, we apply variable structure RBF networks model to forecast total energy in China over a period of 1980 to 2009. The variable structure RBF networks model is used for on-line prediction, for the details, refer to Section 3.3. The data from 2005 to 2009 are used to test predictive precision.

The predictive values are shown in Fig.1 and the mean absolute percentage error (MAPE) of dynamic RBF model and trivial model are 2.29% and 5.31% respectively. The trivial model simply involves using the current value of the option as the next prediction, namely $\hat{y}_{t+1} = y_t$. One can draw several conclusions from the results as follows:

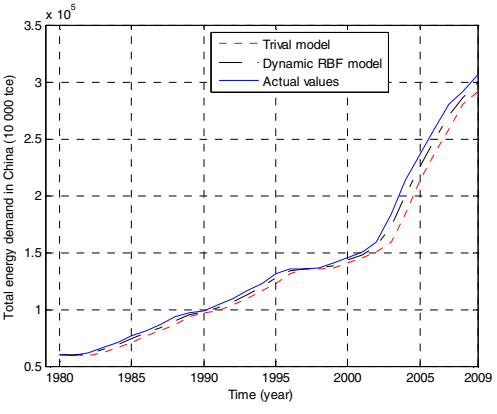


Fig. 1. Prediction of total energy demand in China

(i) The variable structure RBF networks model is applied to forecast energy demand in China and the results show that it can predict the trend efficiently, refer to Fig.1.

(ii) According to MAPE, it can be seen that the forecasting performance of trivial model is worst and its MAPE is 5.31%, while the variable structure RBF networks model has better forecasting performance and its MAPE is only 2.29%.

5 Conclusions

Due to the total energy demand in China shows obviously nonlinearity, a variable structure RBF networks model is proposed in this paper, which has three outstanding advantages as follows: (i) Variable structure RBF networks (i.e. k_t, g_t, q_t vary over time) is adopted to model the nonlinear time series; (ii) Select the input order of RBF networks on-line; (iii) Predict time series on-line based on variable structure RBF networks model using SMC method. The experimental results of total energy demand in China indicate that the variable structure RBF networks model proposed in this paper is effective.

References

1. Chang, C.C.: A multivariate causality test of carbon dioxide emissions, energy consumption and economic growth in China. Applied Energy 87, 3533–3537 (2010) (in Chinese)
2. Wang, J., Bao, Q.: Energy demand forecasting model in China based on wavelet-neural network. Journal of Systems Science & Mathematical Sciences 29(11), 1542–1551 (2009) (in Chinese)
3. Chon, K.H., Richard, J.C.: Linear and Nonlinear ARMA model Parameter Estimation Using an Artificial Neural Network. IEEE Transactions on Biomedical Engineering 44(3), 168–174 (1997)
4. de Freitas, N., Andrieu, C., Hojen-Sorensen, P., et al.: Sequential Monte Carlo Methods for Neural Networks. In: Doucet, A., De Freitas, J.F.G., Gordon, N.J. (eds.) Sequential Monte Carlo Methods in Practice. Springer, New York (2001)
5. Holmes, C., Mallick, B.K.: Bayesian radial basis function of variable dimension. Neural Computation 10(5), 1217–1233 (1998)
6. Doucet, A., Freitas, J.F.G., Murphy, K., Russell, S.: Rao-Blackwellised particle filtering for dynamic Bayesian networks. In: Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, pp. 176–183 (2000)
7. Avrulampalam, S., Maskell, S., Gordon, N., Clapp, T.: A Tutorial on Particle Filters for On-line Non-linear/ Non-Gaussian Bayesian Tracking. IEEE Transactions on Signal Processing 50, 174–188 (2002)
8. Kitagawa, G.: Monte Carlo Filter and Smoother for Non-Gaussian Non-linear State Space Models. Journal of Computational and Graphical Statistics 5(1), 1–25 (1996)
9. Andrieu, C., Freitas, J.F.G., Doucet, A.: Sequential Bayesian Estimation and Model Selection Applied to Neural Networks. Technical report CUED/F-INFENG/TR 341, Department of Engineering. Cambridge University (1999)

A Semantic Service Composition Method Based on Fuzzy Colored Petri Net

Zhe Ma and Kerong Ben

Department of Computer Engineering, Naval University of Engineering
Wuhan, China
mazhe1214@gmail.com, benkerong08@21cn.com

Abstract. As a new application model for decentralized computing, Web services composition is becoming an effective mechanism for the services integration on the Web. For the disadvantage of existing Web service composition methods can only reasoning deterministically, this paper presents a novel semantic Web service composition method with fuzzy colored Petri net. We give the algebra definition and FCPN model of service composition basic structure. On this basis, two-way service composition reasoning algorithm is proposed and experiment shows method effectiveness.

Keywords: Semantic service; fuzzy colored Petri net; service composition.

1 Introduction

Under the service-oriented architecture, Web service composition becomes the main form to realize the business process. However, as Web service doesn't include the semantic information which can be read by computers, how to realize the automatic composition of Web service becomes the big problem in service computing. According to the basis of judging the service's compositional feature, the semantic Web service composition methods can be divided into two kinds: one is judging the service's compositional feature by using the hierarchy relationship of the domain ontology concepts[1-3]. The other according to the behavior of interaction, the composition behavior of sequence, choice, switch can be described through the interaction and cooperation among services[4-6].The first kind of methods describes service as function interface which has the parameter of input and output, but it is difficult to describe the state transformation of service's execution and interaction between services; the other kind focuses on composition process, the model of which based on the model theory tools involves. It judges service composition according to the syntactic match and model checking, it does not make the full use of the semantic information which provides by OWL-S.

Petri net has the formal definition of asynchronous and concurrent, intuitive graphical representation and flexible analysis methods[7]. There is a large number of research works to use Petri net to Web service's formal description, modeling and verification. Valero focuses attention on the development of a methodology for the design and validation of composite Web Services using WS-CDL as the language for describing Web services interactions and Petri net as a formalism that allows to

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 396–404, 2011.

simulate and validate the described systems[8]. YANG primarily extended the Petri net's tokens, which been expressed as the parameters of atomic services, as a model for Web services. And the improved Web service composition method was used to compose the atomic services, making them flowing in the Petri net in accordance with the relationship between the composition of atomic services, and verify the correctness of the service[9]. But in essence, the core of service composition methods is service interactions, and the input and output has been decided when the system model was built. The traditional Petri net does not have enough ability to describe the dynamic characteristics of the service composition implementation. Therefore, this paper proposes a semantic Web service composition method based on fuzzy colored Petri net (FCPN), which uses color function to describe service input/output and reliability of service execution[10], uses reasoning ability of fuzzy Petri net to composite Web services[11].

2 Semantic Service Composition Model Based on FCPN

Semantic Service process is made up by many Web service operations and the control and data flow among them. Every Web service has the attributes such as name, precondition, input data, output result and so on. The control flow decides the sequences of Web service operation and the data flow reflects interaction among the Web service. From the view of organization and interaction, semantic Web service is a set of network operations which uses the standard protocol and interface to do the interaction between applications. Therefore, we define the semantic Web service as a fuzzy colored Petri net structure SWS-FCPN with input and output. The structure uses the color set to describe the quality of service and fuzzy reasoning rule to evaluate the availability of service.

Definition 1 Fuzzy Colored Petri Net Model of Semantic Web Service (SWS-FCPN)

A Semantic Web Service S is defined as $FCPN = (\sum, P, T, A, N, C, I, \alpha, \beta, \gamma, w)$, which means:

1) Color set $\sum = \{data, service\_state\}$, in which $data$ is the value of service input or output, it's usually shows as a string. $service\_state$ is the state of service, and shows as the service reliability with the special input;

2) Place set P is the input and output of the service;

3) Transition set T shows the operation process of atomic service. It's a executive process which is equivalent to a fuzzy reasoning rule's execution;

4) Arc set A and node function N defines the flow relations among services;

5) Colored function $C = \{C(IN) = data, C(OUT) =< data, \quad service\_state >\}$ is the input and output states of service;

6) Initialization function I initializes the state of the net model;

7) Transition threshold function α shows the precondition of the executing service;

8) Place confidence function β shows the confidence of input and output of service;

9) Transition confidence function γ shows the probability of not leading to failure when the service executed;

10) Arc weights function w shows the degree of influence to the service execution by the service input and output.

Atomic service is a process of interaction between single services, so atomic service could be mapped to the transition of net model directly and the input and output of service can be mapped to the Token in the place. FCPN model of atomic service is shown as Fig. 1.

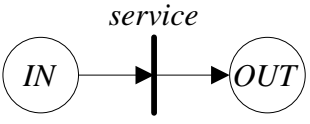


Fig. 1. FCPN model of atom service

In service-oriented business process development, a Web service is considered as a partial order set which has service state and service operation. By service search and interface integration between services, several Web services can be composited to a new service with greater function.

2.1 Basic Composition Pattern

The algebraic definition of basic Web service composition patterns and the corresponding FCPN model is shown as below:

Definition 2 Sequence Composition Operator $\succ$

$ws_1 \succ ws_2$ shows that the executive relationship is sequence when the two service process ws_1 and ws_2 are composited. $\succ$ is sequence composition operator. $ws_1 \succ ws_2$ shows that ws_2 has to execute after ws_1. When Token arrivals output set from input set of ws_1, Token can be transferred into the input set of ws_2 and execute the service if the interface is matched. In the process of BPEL, <sequence> defines the structure of the sequence composition. FCPN model of the sequence composition is shown as Fig. 2.

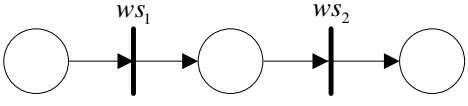


Fig. 2. Sequence pattern

Definition 3 Parallel Composition Operator $\oplus$

$\oplus$ shows that the executive relationship is parallel when the two service process ws_1 and ws_2 are composited. $\oplus$ is parallel composition operator. $ws_1 \oplus ws_2$ shows

that system can be discomposed into several activity branches which can be executed parallel. ws_1 and ws_2 can be executed in the same time. In the process of BPEL, <flow> defines the parallel composition. FCPN model of parallel composition is shown as Fig. 3.

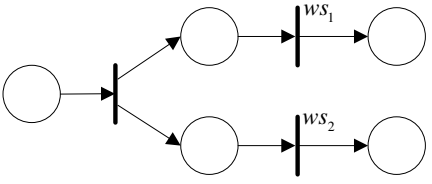


Fig. 3. Parallel pattern

Definition 4 Simply Choice Composition Operator ⊗

⊗ shows that the executive relationship is simply choice when the two service process ws_1 and ws_2 are composited. ⊗ is simply choice combination operator. There must be one and only in ws_1 and ws_2 which can be executed. They cannot be executed at the same time. Similarly, several service processes can composite into simply choice structure. In BPEL process, <switch> defines the simply choice composition. The FCPN model of simply choice composition is shown as Fig. 4.

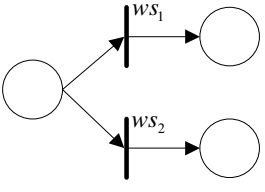


Fig. 4. Simple choice pattern

Definition 5 Iteration Operator ⊙

⊙ shows that ws should be executed repeat till satisfies the condition. In fact, ws is composited by several transitions and places. In the process of BPEL, <while> defines the iteration structure. The FCPN model of iteration structure is shown as Fig. 5.

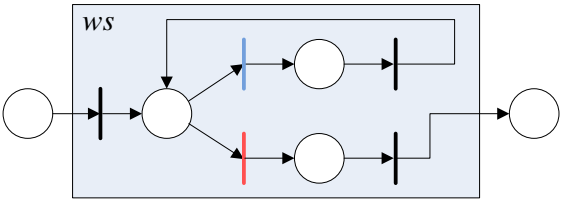


Fig. 5. Iteration pattern

Definition 6 Delayed Choice Composition Operator $\tilde{\otimes}$

$\tilde{\otimes}$ shows that the executive relationship is delayed choice when the two service process ws_1 and ws_2 are composited. $\tilde{\otimes}$ is delayed choice composition operator. Delayed choice is similar to simply choice. However, simply choice must choose a branch to execute when arrive the branch point; while delayed choice has to wait the choice information when arrive the branch point until satisfies the condition. In the process of BPEL, <pick> defines the delayed choice structure. The FCPN model of delayed choice structure is shown as Fig. 6.

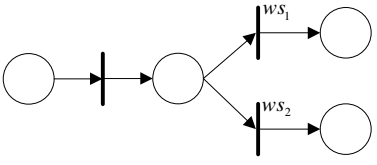


Fig. 6. Delayed choice pattern

Definition 7 Synchronization Operator $\tilde{\oplus}$

$\tilde{\oplus}$ shows that the executive relationship is Synchronization when the two service process ws_1 and ws_2 are composited. $\tilde{\oplus}$ is synchronization operator. Synchronization is the embedded structure of parallel. In the relationship of parallel composition, the synchronized service process is independent with each other. Synchronization can assign the sequence of concurrent services. $ws_1 \tilde{\oplus} ws_2$ means that ws_1 and ws_2 are parallel services, but ws_2 must be executed after ws_1. In the process of BPEL, <link> defines the synchronization structure. The FCPN model of synchronization composition structure is shown as Fig. 7.

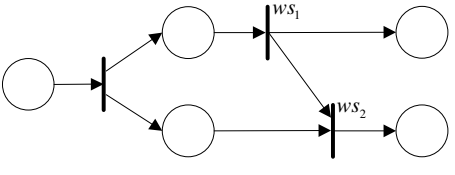


Fig. 7. Synchronization pattern

A complex service composition process can be connected and shown by the composition operators as below and simplified by the algebraic features of the composition operators.

2.2 Semantic Web Service Composition Fuzzy Reasoning

In the period of service construction and arrangement, a service composition model can be built based on FCPN, according to the independence relation of the input and output

between Web services. It's also can be considered as abstractive description of the software. Although the construction of FCPN model is complex and the network scale is large, it can be used repeatedly after constructed once. In the period of system execution, relative semantic Web services will be searched in the FCPN network according to the requirements of users and corresponding function will be implemented according the process structure.

The operation process of FCPN is a reasoning process according to the production rules. The basic form of FCPN reasoning rules is shown as following formula which consists with two propositions d_j and d_k, which can be correspond to the place p_j and p_k.

$$R_i : IF \ \ d_j \ \ THEN \ \ d_k \quad (CF = \mu_i) \tag{1}$$

The confidence of transition is μ_i. If the fidelity of proposition d_j is β_j, then after transition the fidelity of proposition d_k is $\beta_k = \beta_j \times \mu_i$. It is shown as Fig. 8.

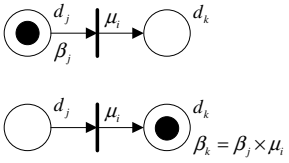


Fig. 8. FPN reasoning process

The process of reasoning can adopt different strategies according to different applications. We use forward and backward reasoning and simplify the reasoning result by using the two-way reasoning. Firstly, the input parameters and output parameters of users' needs are matched with the services which have already in the Web service library, get the place set P_{IN} which matches input parameters and P_{OUT} matched with output parameters. According to the backward reasoning, gradually finds and saves all the places and transitions which can deduce. Then set from the input place set P_{IN} deduces P_{OUT} according to forward reasoning. Saves all the transition sets and place sets which involve in the process of reasoning. Finally, generate the service composition model which can satisfy the needs and realizes the auto semantic service composition.

3 Case Study

Consider the scene of a stock exchange, a complete process of stock exchange consists four basic activities: stock query, stock purchase, stock sale and inform clients. For the sake of simplicity, suppose that there are three stocks: Stock A, Stock B and stock C, the client's accounts are Stock Account and Money Account, the ways to inform clients are mail and short message. The FCPN model is shown as Fig. 9.

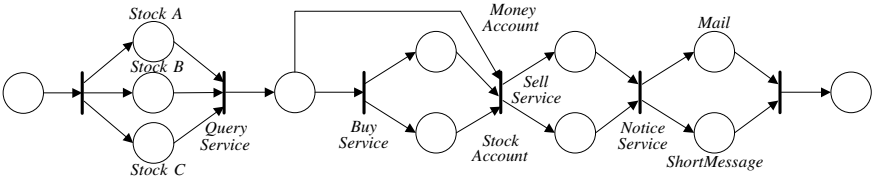


Fig. 9. FCPN model of stock bargaining

In Fig. 9, the four basic services, stock query, stock purchase, stock sale and inform clients are shown by single transition. The corresponding fuzzy reasoning rules are shown as table 1.

If the client's need is selling Stock B and receiving short message, the input parameters and expected output parameters will be p5 and p12. According to algorithm, do the forward and backward reasoning.

Table 1. Fuzzy reasoning rules of stock bargaining

Service Name	Reasoning Rules
Query Service	*IF StockType THEN StockPrice*
	$CF_1 = 0.8$
Buy Service	*IF StockType AND Amount THEN*
	MoneyAccount AND StockAccount
	$CF_2 = 0.98$
Sell Service	*IF StockType AND Amount THEN*
	MoneyAccount AND StockAccount
	$CF_3 = 0.98$
Notice Service	*IF BargainingInformation THEN*
	Mail OR ShortMessage
	$CF_4 = 0.7$

If the client's need is selling Stock B and receiving short message, the input parameters and expected output parameters will be p5 and p12. According to algorithm, do the forward and backward reasoning. The FCPN model of final service composition result is shown as Fig. 10.

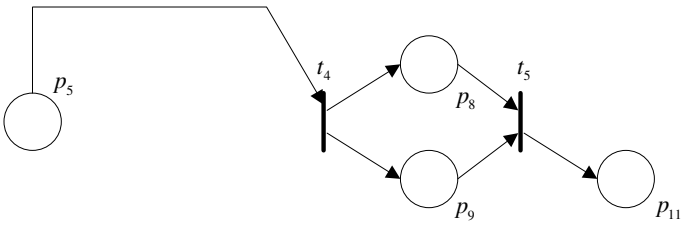


Fig. 10. Service composition result

The confidence of service composition result is $\beta(p_{11}) = \beta(p_5) \times \gamma(t_4) \times \gamma(t_5) = \beta(p_5) \times 0.98 \times 0.7 \quad = 0.686\beta(p_5)$. If the amount of stock purchase which input by the client is more than the amount of stock that he owns, the confidence of this input will be $\beta(p_5) = 0$, else $\beta(p_5) = 1$. It can be seen that when the amount of stock which was input by the client is over the limitation, the reliability of service composition result that he finish the exchange and receive the short message will be 0.

4 Conclusion

In this paper we analyzed the disadvantages in the abilities of semantic expression and system description of the existing Web service composition method and put forward a semantic service composition method based on fuzzy colored Petri net which combining the features of Petri Net and fuzzy Petri net. We also propose the algebraic definition of basic structure of service composition patterns and corresponding FCPN model. Based on this method, a two-way reasoning algorithm of service composition was proposed and the effectiveness of was explained by giving some cases. As the semantic service composition method of FCPN kept the formal description and verification capability to the service interaction of the classic Petri Net, increased the description of the reliability to the service operation to the different input. The service composition was accomplished by fuzzy reasoning rule. The method overcomes the disadvantage of lack reasoning support ability of which described by traditional semantic Web service composition based on OWL-S.

References

1. Bellur, U., Kulkarni, R.: Improved Matchmaking Algorithm for Semantic Web Services Based on Bipartite Graph Matching. In: Proceedings of the, IEEE International Conference on Web Services, pp. 86–93. IEEE Computer Society Press, Salt Lake City (2007)
2. Srivastava, U., Munagala, K.: Query Optimization over Web Services. In: Proceedings of the 2006 International Conference on Very Large Data Bases, pp. 355–366. ACM Press, Seoul (2006)

3. Roman, D., Kifer, M.: Reasoning about the Behavior of Semantic Web Services with Concurrent Transaction Logic. In: Proceedings of the 2007 International Conference on Very Large Data Bases, pp. 627–638. ACM Press, Vienna (2007)
4. Salaun, G., Bordeaux, L., Schaerf, M.: Describing and Reasoning on Web Services using Process Algebra. International Journal of Business Process Integration and Management 1(2), 116–128 (2006)
5. Keum, C., Kang, S., Ko, I.: Generating Test Cases for Web Services Using Extended Finite State Machine. In: Proceedings of International Federation for Information Processing, pp. 103–117. Springer, Santiago (2006)
6. Berrick, W., Leptoukh, G., Giovanni, F.D.: A Web Service Workflow-Based Data Visualization and Analysis System. IEEE Transactions on Geoscience and Remote Sensing 47(1), 106–113 (2009)
7. Petri, C.A.: Kommunikation mit Automaten. University of Bonn, Bonn (1962)
8. Valero, V., Cambronero, E., Macia, H.: A Petri Net Approach for the Design and Analysis of Web Services Choreographies. Journal of Logic and Algebraic Programming 78(5), 359–380 (2009)
9. Yang, B., Song, L.-S.: Web Service Automatic Composition Based on Extended Token of Petri Net. Applied Mechanics and Materials (20-23), 220–225 (2010)
10. Jensen, K.: Coloured Petri Nets: Basic Concepts, Analysis Methods, and Practical Use, pp. 65–85. Springer, Heidelberg (1997)
11. Fu, R., Dong, W.-Q., Yang, G.: Fuzzy Petri Net-Based Optimized Semantic Web Service Composition. In: Proceedings of the 2008 International Conference on Grid and Cooperative Computing, pp. 496–502. IEEE Computer Society Press, Shenzhen (2008)

Interactive and Collaborative Games Promoting Metacognition for Science and Engineering Design

Ying Tang[1] , Sachine Shetty[2], John Henry[3],
Kauser Jahan[1], and Samuel Hargrove[2]

[1] College of Engineering, Rowan University, 201 Mullica Hill Rd., Glassboro, NJ 08028
{tang,jahan}@rowan.edu
[2] College of Engineering, Tennessee State University, 3500 John A Merritt Blvd,
Nashville, TN 37209
{sshetty,skhargrove}@Tnstate.edu
[3] Education Information Resource Center, 107 Gilbreth Parkway,
Suite 200 Mullica Hill, NJ 08062
jhenry@eirc.org

Abstract. The focus of this project is to design and implement a virtual reality game system that infuses cyberinfrastructure (CI) learning experiences into the engineering classrooms to promote metacognition for science and engineering design in context. Using city infrastructure as the theme and engineers solving real-life problems as the scenes, the games of future sustainable city design engage students, particularly beginning science and engineering students, in CI-enhanced and -enabled science and engineering discovery. The CI features, the metacognitive strategies, the context-oriented approaches as well as their seamless integration in the game are presented through a detailed game module.

Keywords: Virtual Reality Game, Metacognition, Science and Engineering Design.

1 Introduction

The radical and transformative technological revolution has resulted in fundamentally new ways of science and engineering practice. This paradigm shift has a significant impact on the skills needed for a diverse science and engineering workforce that is capable of designing and deploying cyber-based systems, tools and services. However, our education has not kept pace with this evolution. In fact, there is a crucial need to bring cyberinfrastructure (CI) learning experiences into classrooms of higher education. Two key CI based technologies which have tremendous impact on education and training are 1) networked computing technologies, and 2) virtual learning environments, including games, simulations and modeling. The former, such as networked desktops and laptops, enables new forms of collaborative learning that responds or gathers information from the environment in a time- and location-aware manner to meet different learners' requirements. There is a growing use of virtual environments for work, play, entertainment, and scientific inquiry. Games have come a long way to be much more than visualization. They are interactions within

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 405–412, 2011.

immersive digital worlds that promote learning through authentic and engaging play. Simulations and models help provide insights into scientific phenomena, making difficult abstract concepts and large data sets accessible in ways that are more visual, interactive, and concrete. As such, infusing virtual reality (VR) games with simulations and models into a classroom setting becomes essential.

Living through the hurricane Katrina and its aftermath and reflecting on these experiences from the technical and humanist standpoints has led us realize the importance of bringing the perspectives of humanities and social sciences into design education [1]. Research indicated that such design education should start earlier to prospective and beginning science and engineering students, encouraging them to optimize their design activities not only in technical aspects, but also in social and environmental standpoints [2].

Motivated by these general remarks, this project, as a collaboration between Rowan University, Tennessee State University (TSU), and Education Information Resource Center, develops and implements a VR game system that provides CI learning experiences with visualization, collaboration, and simulation tools to promote metacognition for science and engineering design in context. With minimal or no modifications to the existing course materials, project-based participatory games, demonstrating science and engineering principles in real engineering problem-solving, are designed to replace the traditional laboratory setting of the introductory engineering courses.

2 Metacognitive Interventions

Awareness and monitoring of one's learning processes with a repertoire of strategies that they apply when and as required by different learning circumstances are increasingly recognized as critical for successful learners. Such awareness and monitoring processes are often refer to as metacognition –" the processes in which the individual carefully considers thoughts in problem solving situations through the strategies of self-planning, self-monitoring, self-regulating, self-questioning, self-reflecting, and or self-reviewing" [3]. Four important metacognitive interventions, as detailed below, are carefully designed into the interactive game activities.

- Road Map training – Learning roadmap provides study guides that endow pupil with the capability to find the relevant information and to capture the concepts in the study materials [4]. In this project, a study guide, as exemplified in Fig. 1, is a set of suggestions designed to lead students through a problem solving process by directing attention to the key ideas and suggesting the application of skills needed to solve the problem successfully.
- What I **K**now-What I **W**ant to Know- What I have **S**olved (KWS) training – KWS is adapted from a well known reading strategy, What I **K**now-What I **W**ant to Know- What I Have **L**earned (KWL) [5]. It typically provides a 3-column chart structure as exemplified in Fig. 2, to activate students' prior knowledge by recalling what students know about a problem (K), to motivate students to read/think by asking what they want to know (W), and finally to review what part of the problem has been resolved and what is yet to be solved (S).

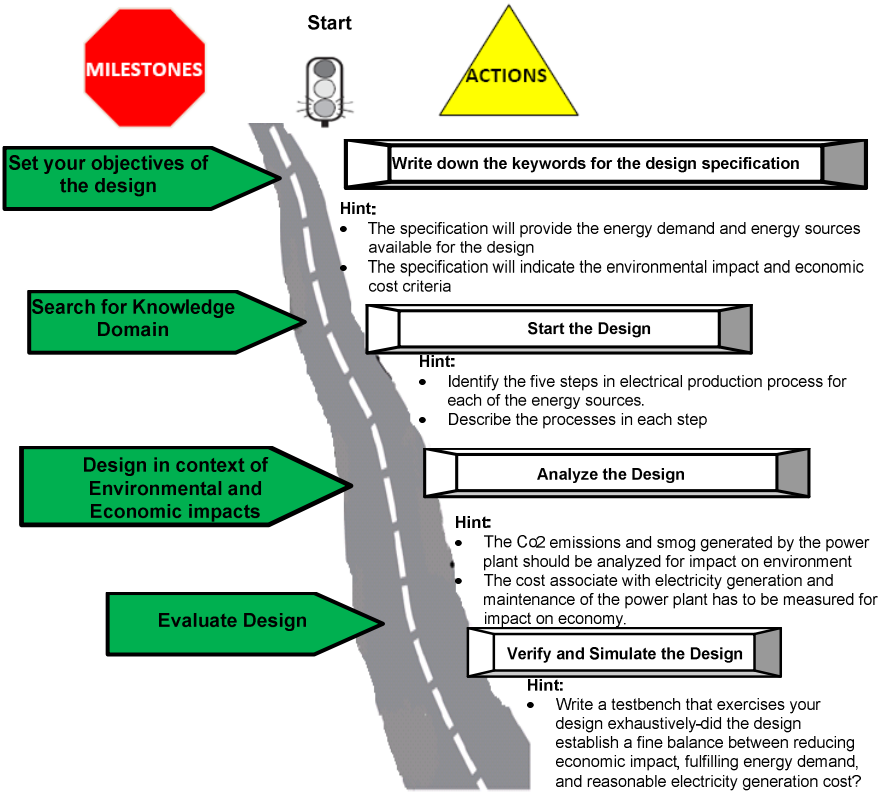


Fig. 1. A sample road map for power plant design

KWS - Power Plant Design

What Do I Already **K**now	What Do I **W**ant To Know	What Have I **S**olved
• Electrical Production: Renewable and Non-renewable Energy Sources • Non-Renewable Energy Sources: Fossil Fuels (Gas, Oil and Coal), • Renewable Energy Sources: Solar, Nuclear, Hydro Electric, and Wind • Steps in Electrical Generation: Mining, Refining, Transporting, Generation, Transmission, etc.	➢ How to assess the environmental impact due to pollution effects on plants and human beings, greenhouse effect, global warming effects? ➢ How to assess the economic impact due to market price of energy sources, government regulations, transportation costs, import of electricity from another region?	✓ Electrical Generation Steps for all energy sources.

Fig. 2. A sample KWS worksheet

- Concept Map – One of the most valuable strategies for improving students' self-reflection and strategic thinking is through a process of creating and modifying a concept map [6]. In this project, the concept mapping activities are used closely with the KWS training. At the beginning of each game, a preliminary concept

map essentially constructed by the game system, as shown in Fig. 3, is provided to students. Points are awarded for the students who identify any concept variables of the preliminary map in the *K* column of their KWS chart. When students see their ideas are, in fact, the basis to a design or problem-solving, their self-efficacy is expected to be increased significantly. Throughout the game, the concept map is revisited regularly as students further explore the problem and obtain more knowledge. While students continue to verify, clarify and expand, or replace each of their initial statements with more accurate information, more unknown becomes known and is moved from the *W* to the *K* column in the KWS chart.

- Think-Aloud-Share-Solve (TA2S) training – As Vygotsky pointed out, learning is an inherently social and cultural rather than individual phenomenon [7]. The interactions among peers produce intellectual synergy of many minds to bear on a problem, and the social stimulation of mutual engagement in a common endeavor. TA2S is a variation of the collaborative learning strategies, Think-Aloud [8] and Pair Problem-solving [9], which is implemented in our game system through online chatting..

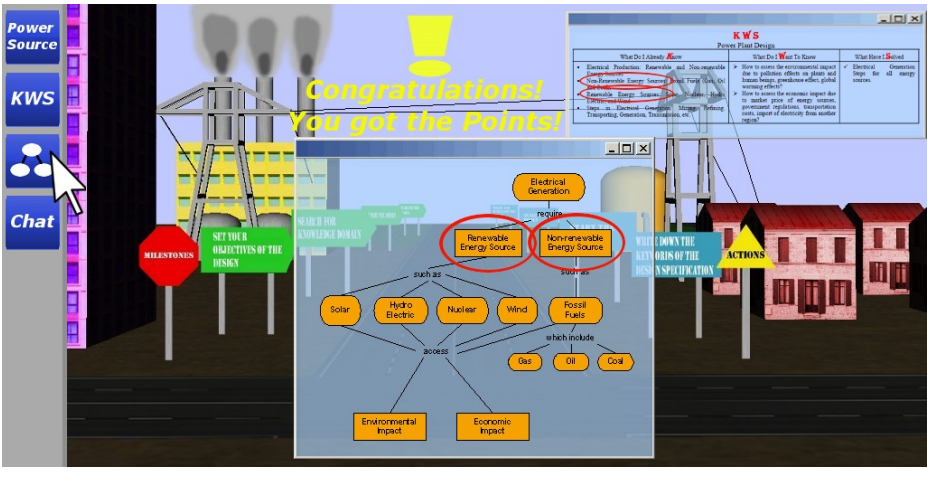


Fig. 3. A sample concept map for power plant design

3 Game Theme and Design

Recent reports have indicated that many STEM classes rely heavily on textbook but weak on examples, so that students are exposed to encyclopedias of fact without ever engaging in the science and engineering process [11]. The perceived dullness or complexity of the material, a lack of concrete applications, and preconceptions among students further make introductory science and engineering classes difficult to students, leading to lower recruiting and retention rates of science/engineering majors. Thus, it is crucial to design a fun learning environment that engages students in the exploration of real science and engineering applications where strategic, constructive,

and big-picture thinking and problem solving are promoted. Considering the entertaining, experiential and problem-based nature of games, the proposed VR game system then fits perfectly into this scope to effectively bring to students knowledge and skills that might otherwise be just bullet points on slides.

With the *Future Sustainable City* as a broader context, and the *City Infrastructure* as the theme, our game system consists of several courses at Rowan and TSU have an important laboratory component, where we implement our proposed games as a replacement to the traditional lab experiments. The game is a group assignment that requires each student in a group to log in the system with a security password to play it. While his/her login ID becomes visible to other players in the group, group discussion on the problem/solutions is then enabled through cyber-chatting, where group members are not necessarily present at the same place. The discussions as well as each player's actions will be recorded in the system only accessible to instructors and researchers, providing a good resource to analyze student performance and game effectiveness in promoting learning.

In this section, this educational process is exemplified in a game module where the seamless integration of fun, metacognitive interventions and engineering problem-solving in a well-balanced engagement and learning process is showcased.

Game: Power Ville – Cooking a dinner, heating a house, lighting a street, and running a factory, all of these need power. Energy is thus at the heart of everybody's quality of life. How to generate and use energy that satisfies the increasing energy needs and combats climate changes at the same time then becomes an unprecedented challenge for a sustain-city development. Bring such real science and engineering design problem as well as involved societal and environmental issues into the introductory science and engineering courses is the core of this game project.

Starting with a flash movie similar to "Energy Crossroad" [10], students are introduced to the game with a brief exposure of current energy consumption, their damaging effects on the environment and the global economy, and viable solutions that ease these destructive consequences. A prologue narrative with a design problem is then presented to the students (Fig. 5) –"as a decision maker for the city, how would you design an electrical production system which will provide the city with enough power to meet the energy demands?" Meanwhile, a road map of design guidance is provided to the students as shown in Fig. 6. (Please refer to Fig.1 as well for the detailed milestones and actions). While students walk through the map and advance from one stage to another, they are provided clues related to the KWS datasheet. In particular, these clues convey the presence of renewable and non-renewable energy sources, five steps to implement the electrical production of any chosen energy source, and its environmental impact and economic costs involved. While the students complete their initial KWS, a preliminary concept map is then given as depicted in Fig. 3. As stated earlier, students will be awarded if they identify any concept variables of the preliminary map in the *K* column of their KWS chart.

One of the critical design steps involved in the game is to implement the five steps required to design an electrical production system: Mining, Refining, Transporting, Generation, and Transmission, and to analyze their environment impact and economic costs. The crux of a sustainable city is to create the smallest ecological footprint, and to produce the lowest quantity of pollution possible. For any chosen energy source, the CO2 emissions and smog-causing pollutants to be created might be different.

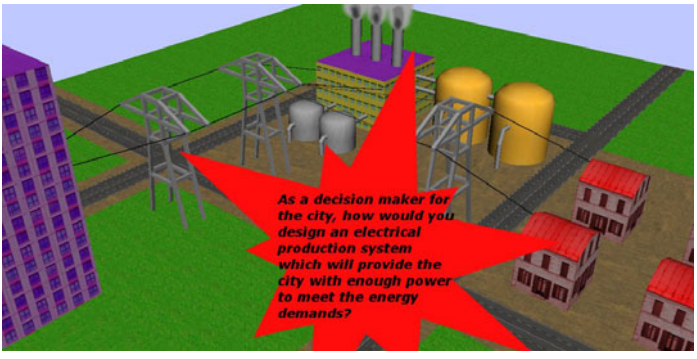


Fig. 5. The game "Power Ville"

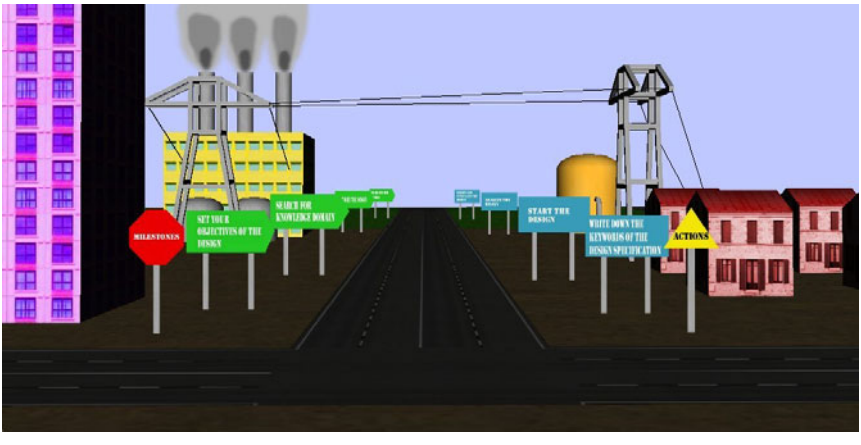


Fig. 6. The road map

As we know, increased instances of smog will raise levels of illness among citizens and keep them from work (which costs the player money). And increased carbon emissions could result in floods, droughts, powerful storms, etc which will also cost money to rebuild the city. So the player has to manage the power resource such that they recognize the costs and trade-offs of their choices. Our game will be designed to enable the player visually viewing such impact with a chosen energy source. As exemplified in Fig. 7, when the player chooses nuclear source and provides necessary parameters (e.g., the total energy demands of the city), an interactive simulator displays the nuclear generation phases and its potential impact on the city due to the estimated amount of CO2 emissions and smog-causing pollutants. To assess the economic costs of a design, the game involves students in computer programming activities. In particular, the students will develop C++ or JAVA code to implement a given optimization model that ensures their design to meet a certain electricity demand at the lowest cost. The game becomes even fun when students play with their on-line group members where ideas and knowledge are shared as shown in Fig. 8. It is in such communication that much of the learning occurs and an optimal solution results from.

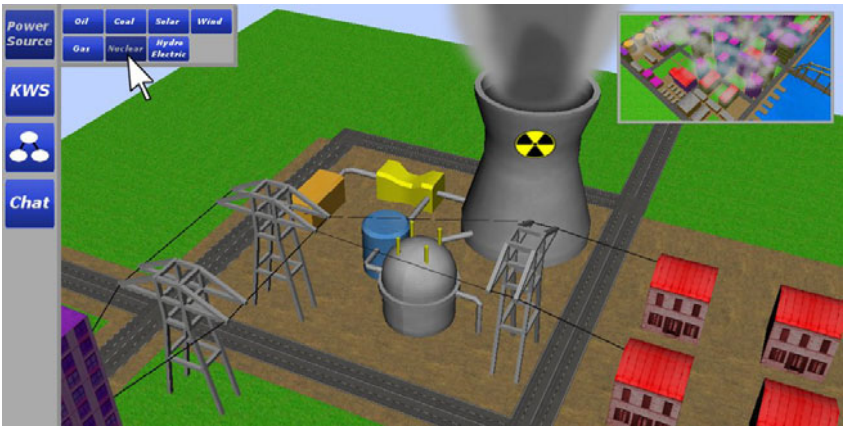


Fig. 7. The interactive simulator for the nuclear generation phases and their emission impact on the city

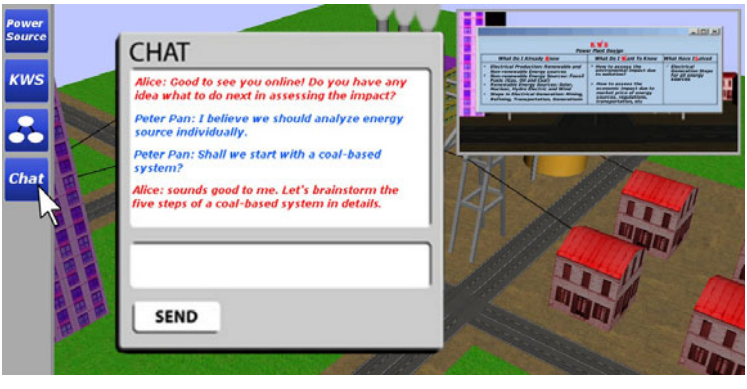


Fig. 8. The player is chatting with her group member online

4 Conclusion

Series of games with the focus on metacognitive and problem-solving strategies are under development and being implemented in four targeted courses at Rowan and TSU, respectively. The learning materials as well as their organization optimize student learning and help them develop a strong foundation to carry over for subsequent coursework

Acknowledgments

This work is supported in part by the National Science Foundation grant #OCI-1041306 and #EEC-0935089.

References

1. Lima, M.: Engineering education in the wake of hurricane Katrina. Journal of Biological Engineering 1, 6 (2007)
2. Yasuhara, K., Morozov, A., Kilgore, D., Atman, C., Loucks-Jaret, C.: Considering life cycle during design: a longitudinal study of engineering undergraduates. In: Proceedings of American Society for Engineering Education Annual Conference, AC 2009, p. 1728 (2009)
3. Hyde, A., Bizar, M.: Thinking in context. Longman, White Plains (1989)
4. Oliveira, M., Serrano, J.A.: Learning roadmap studio: new approaches and strategies for efficient learning and training processes, http://www.elearningeuropa.info/files/media/media16938.pdf
5. Ogle, D.M.: KWL in action: Secondary teachers find applications that work. In: Disher, E.K., Bean, T.W., Readence, J.E., Moore, D.W. (eds.) Reading in the Content Areas: Improving Classroom Instruction, 3rd edn., pp. 270–281. Kendall-Hunt, Dubuque (1992)
6. Cassata, A.E., French, L.: Using Concept Mapping to Facilitate Metacognitive Control in Preschool, Children. In: Cañas, A.J., Novak, J.D. (eds.) Concept Maps: Theory, Methodology, Technology. Proceedings of the Second International Conference on Concept Mapping, Universidad de Costa Rica, San Jose (2006)
7. Vygotsky, L.S.: Mind in society: The development of higher psychological processes. Harvard University Press, Cambridge (1978)
8. Blakey, E., Spence, S.: Developing Metacognition. Educational Resource Information Center (ERIC), http://www.education.com/reference/article/Ref_Dev_Metacognition/
9. Narode, R.B.: Pair Problem-Solving and Metacognition in Remedial College Mathematics. Technical Report, http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1c/53/3b.pdf
10. Energy Crossroad – a documentary film, http://www.energyxroads.com/about.html
11. Herbert, B.E.: The role of scaffolding student metacognition in developing mental models of complex, Earth and environmental systems. Paper Presented at DFG-NSF International Workshops on Research and Development in Mathematics and Science Education, Washington, DC (November 2003)

A Personalized Mobile Application Using Location Based Service

Lovely Chatterjee, Saswati Mukherjee, and Matangini Chattopadhyay

School of Education Technology,
Jadavpur University
{lovely.chatterjee,chttpdhy}@yahoo.com, sash_cal@rediffmail.com

Abstract. Technology has transformed the cell phone from being a mere phone to become an indispensable medium for contact, information and entertainment. One of the most common features of mobile communication is *Location Management*, which keeps track of the locations of the mobile users who leave the home location and roam into foreign network. Location Based Service (LBS) plays an important role in the rapidly evolving world of mobile applications. In this paper, a novel personalized application using location based service has been proposed and implemented. When the mobile user is away from its home location, he/she can launch that application which automatically sends SMS to those mobile numbers present in the "contacts" applicable to visitor location i.e., friends and relatives in the visitor location.

Keywords: Location Management, Record Store Management, Wireless Messaging.

1 Introduction

A wireless network is an emerging technology that has become increasingly popular in the computing era. Mobile wireless networks allow users to access information and services anywhere, anytime regardless of their geographic positions. There are currently two variations of mobile wireless networks: - infrastructure based networks and infrastructure less networks. One of the most common features of mobile communication is *Location Management,* which keeps track of the locations of the mobile users who leave the home location and roam into foreign network. Various applications based on LBS have become useful to provide community service. The present work is a personalized mobile application that is highly effective for those who want to automatically get connected to friends and relatives present in the visitor location. The proposed application contains three components - Location Management, Record Store Management and Wireless Messaging.

Location Management. A GSM or UMTS network, like all cellular networks, is a radio network of individual cells, known as base stations. Each base station covers a small geographical area, which is part of a uniquely identified location area. A group of base stations is given a name called, a location area, or a routing area. The location update procedure allows a mobile device to inform the cellular network, whenever it

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 413–419, 2011.

moves from one location to the next. Mobile devices are responsible for detecting location area codes. When a mobile finds that the location area code is different from its last update, it performs another update by sending to the network, a location update request, together with its previous location, and its Temporary Mobile Subscriber Identity (TMSI).

Record Store Management. In MIDP persistent storage is centered on the record store, which is a small database. Record stores are represented by instances of javax.microedition.rms.RecordStore. The scope of a record store can be limited to a single MIDlet or shared between MIDlets. Record stores are identified by a name. Within a MIDlet suite the names of the recordstores must be unique. Each record in a record store is an array of bytes and has a unique integer identifier.

Wireless Messaging. The Wireless Messaging API (WMA) is an optional package for J2ME that provides platform-independent access to wireless communication resources like Short Message Service (SMS). It thus supports Java 2 Platform, Mobile Edition (J2ME) applications targeted at cell phones and other devices that can send and receive wireless messages. All the WMA components are contained in a single package, *javax.wireless.messaging*, which defines all the interfaces required for sending and receiving wireless messages, both binary and text.

The paper is organized as follows. Section 2 discusses about already published works on the development of location-based mobile applications. Section 3 describes the considered scenario and the design of the proposed application. In Section 4, the implementation and the screenshots of the simulators are presented. Finally, Section 5 draws the conclusions on the work presented in this paper.

2 Related Work

There are many mobile application based on location management so far. We have mainly concentrated on finding the exact location information (Latitude and Longitude information) of a mobile device and performing certain task depending on the location information of the device and also wireless messaging.

There are many location-based applications in real world. Many location management based strategies are proposed in [1], [2], [3], [4], [5], [6]-[9]. In [1] there is a scheme to update the location depending on the mobility pattern of the mobile device whereas [2] and [3] propose a scheme which reduces HLR traffic while changing locations. In [2] authors brought about network overload of the HLR when a mobile user moves back and forth between adjacent location areas. It can solve either of the location updating rate or location updating cost. But the scheme in [3] provided low location updating traffic and also was applicable to manage both the location updating cost and rate. On the other side [4] and [5] introduce a proposal of updating locations through reminders. There was a model in [4] which delivered location based information to the mobile user depending on the location and time. The scheme in [5] could minimize the overall transmission costs of produced messages by intelligently adjusting the ratio between updating messages and forwarding pointers. [6]-[9] propose different types of schemes in location management to create a mobility management network by minimizing costs. In [10]-[15], [16], [17], [18] there are

different type of developments of mobile applications. In these papers, the authors wanted to use the mobile device not only to make call or receive calls but also as a mobile diagnosis system, a product identification device, a tourists guide and also, a great interaction system among human beings by the methods of touching pointing and scanning.

By going through all these papers, an attempt has been made to design a system that informs the presence of a mobile user to his/her relatives and friends who reside on the mobile user's visitor location.

3 Proposed Work

The proposed application first finds location of the mobile user and then sends an SMS to the contacts pertaining to the same location. Fig. 1 shows the architecture of the proposed application. Firstly, the mobile user launches the application that finds the location information in latitude longitude form. It then retrieves the location name. After finding the location information, it will check its database (Record Store) in search of those numbers that have the same location. When the numbers are retrieved, the application will send a message to those numbers at a time.

When a user changes its position, the location of its mobile device changes. Then the user has to launch the application. The internal structure of the application consists of three classes. The classes are for finding location, storing numbers and sending messages. The names of the classes are "Location", "RecordStore" and "SMSSend". The application is developed using J2ME.

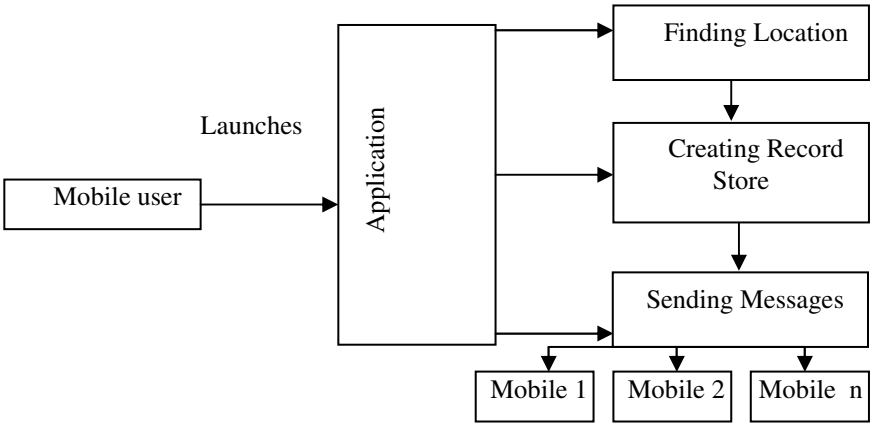


Fig. 1. Block Diagram of the Proposed System

The class, Location, finds the latitude and longitude with the help of getLatitude() and getLongitude() methods respectively. It also finds the location name with help of getLocationName() method. After retrieving the location name, it calls the RecordStore class and forwards the location name to that class. The following code segment is used to get latitude and longitude of a mobile device. The variables called

"*lat*" and "*lon*" are used to store the latitude and longitude obtained using *getLatitude()* and *getLongitude()* methods of the object *c* of *coordinate* class. The variable, *coord,* is used to store the coordinate in the form of latitude and longitude.

```
double lat = c.getLatitude();
double lon = c.getLongitude();
String coord = "("+lat+","+lon+")";
```

The RecordStore class now has the location name of the user. It then checks its database, which keeps the contacts' names, locations and numbers. The following code segment is used to get the contact location name, contacts name and their numbers.

```
custo_id[i]=inputDataStream.readInt();
custo_name[i]=inputDataStream.readUTF();
custo_loc[i]=inputDataStream.readUTF();
String num1=Integer.toString(inputDataStream.readInt();
```

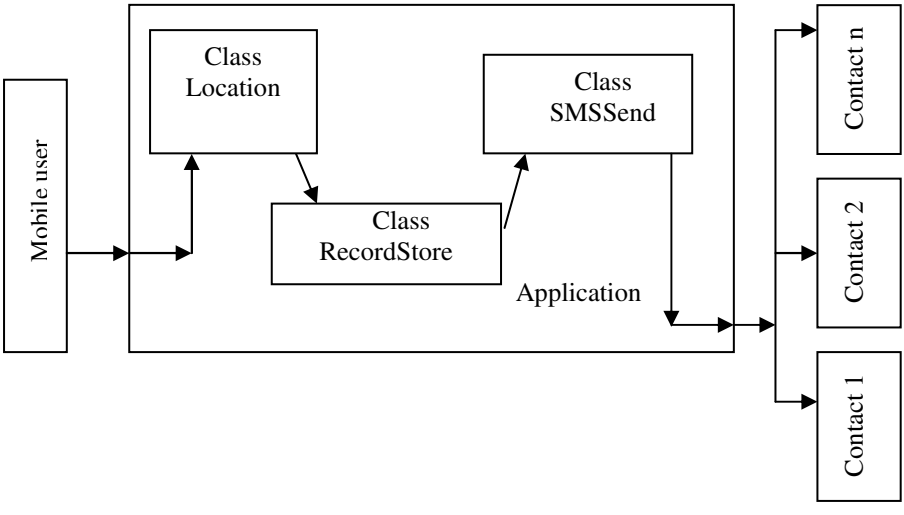


Fig. 2. Classes and their interactions

While checking, if any location matches with that of the user then it stores those corresponding number of those contacts in a vector. After that it dispatches the vector to SMSSend class. It is the responsibility of the SMSSend class to send a hello message to those numbers present in the vector.

```
if(custo_loc[i].equals(s2)){
v.insertElementAt(num1,n);
n++;
}
```

Thus, using the proposed application anyone may inform his/her presence to his or her relatives or friends. Fig. 2 shows various classes and their interactions in the application.

4 Implementation

The application is implemented using J2ME. Sun Microsystems Wireless Toolkit 2.5.2 has been used. The application is working in WTK simulator.

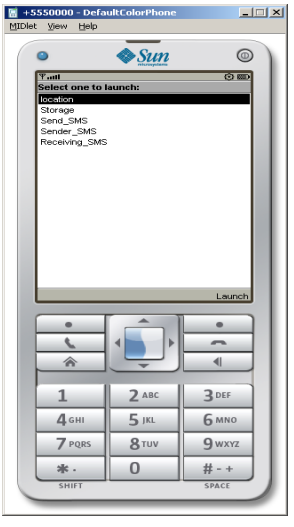


Fig. 3. After running the application

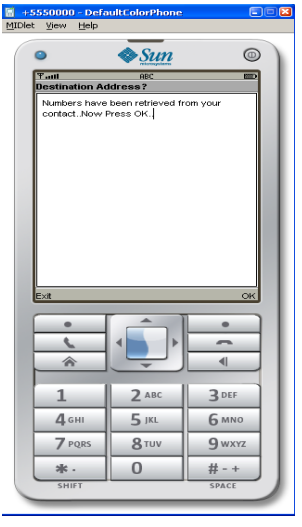


Fig. 4. Retrieving numbers

Fig. 5. Composing message

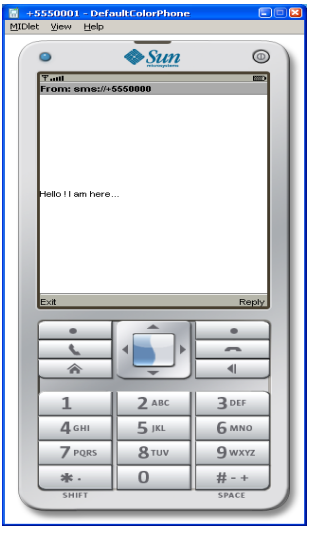


Fig. 6. Receiving the message

When the application is launched by the user (See Fig. 3), it first retrieves the location name and searches it record store for the contacts with the same location name. After finding the numbers (See Fig. 4), it goes to message sending stage where the user can compose message (See Fig. 5).

After composing the message, if user presses the send button, application sends the message to the destinations, which are retrieved.

Then the message will be received by more than one mobile simulator at a time. One of them is shown in Fig. 6.

5 Conclusion and Future Work

In this paper a location management based personalized application has been presented. This application introduces a reminder system to the mobile users. The application sends SMS to the relatives and friends whenever the user is in roaming condition.

This application is well proved in simulators but needs testing in real device. Also getting the location is only possible in device with higher version like E-series or C7 etc. Therefore, the proposed application is worthy of consideration in real cellular systems and definitely, it will work as it has been stated.

References

1. Martinez-Arrue, I., Garcia-Escalle, P., Casares-Giner, V.: Location Management Based on the Mobility Patterns of Mobile Users. In: Cerdà-Alabern, L. (ed.) EuroNGI/EuroFGI 2008. LNCS, vol. 5122, pp. 185–200. Springer, Heidelberg (2008)
2. Lee, D.C., Kim, G.H., Yoo, S.J.: Mobility Grouping Scheme to Reduce HLR Traffic in IMT-2000 Networks. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3215, pp. 110–115. Springer, Heidelberg (2004)
3. Chang, C.C., Lin, J.C., Lin, C.C.: A Novel Location Tracking Scheme for Reducing Location Updating Traffic in a Personal Communication System. Springer Science+Business Media, LLC (2007)
4. Ludford, P.J., Frankowski, D., Reily, K., Wilms, K., Terveen, L.: Because I Carry My Cell Phone Anyway: Functional Location-Based Reminder Applications. In: CHI 2006 Proceedings, April 22–27. ACM, New York (2006); 1-59593-178-3/06/0004
5. Yeh, T.Y., Wang, T.I.: A Ratio-Based Update Scheme for Mobile Agent Location Management. In: Håkansson, A., Nguyen, N.T., Hartung, R.L., Howlett, R.J., Jain, L.C. (eds.) KES-AMSTA 2009. LNCS (LNAI), vol. 5559, pp. 100–109. Springer, Heidelberg (2009)
6. Kang, J., Li, Q., Zhu, H., Wu, W.: Formal Approaches to Location Management in Mobile Communications. In: Janowski, T., Mohanty, H. (eds.) ICDCIT 2010. LNCS, vol. 5966, pp. 243–254. Springer, Heidelberg (2010)
7. Yang, S., Luo, H., Qin, Y., Zhang, H.: Design and Evaluation of DNS as Location Manager for HIP. Springer Science+Business Media, LLC (2008)
8. Xinrong, L.V., Jianxin, L.I.A.O., Xiaomin, Z.H.U., Jiachun, W.U.: Difference-less mobility management network. Front. Electr. Electron. Eng. China 4(1), 24–30 (2009)

9. Li, K.: Cost analysis and minimization of movement-based location management schemes in wireless communication networks: a renewal process approach. Springer Science+Business Media, LLC (2011)
10. Rukzio, E., Broll, G., Leichtenstern, K., Schmidt, A.: Mobile Interaction with the Real World: An Evaluation and Comparison of Physical Mobile Interaction Techniques. In: Schiele, B., Dey, A.K., Gellersen, H., de Ruyter, B., Tscheligi, M., Wichert, R., Aarts, E., Buchmann, A. (eds.) AmI 2007. LNCS, vol. 4794, pp. 1–18. Springer, Heidelberg (2007)
11. Wilson, D., McLoughlin, E., O'Sullivan, D., Bertolotto, M.: Contextual Media Integration and Recommendation for Mobile Medical Diagnosis. In: Wade, V.P., Ashman, H., Smyth, B. (eds.) AH 2006. LNCS, vol. 4018, pp. 333–337. Springer, Heidelberg (2006)
12. Fladenmuller, A., De Silva, R.: The effect of Mobile IP handoffs on the performance of TCP. Ó Baltzer Science Publishers BV (1999)
13. Ala-Kurikka, J., Ohtonen, J., Harjula, E., Ylianttila, M.: Improving Multiple Mobile Application Interaction with Unified Session Management. In: The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2006). IEEE, Los Alamitos (2006)
14. Kulkami, S., Diwan, S., Bansode, N.K.: Device Independent Mobile Application Controller for Remote Administration of a Server over a GPRS Link using a J2ME Cellular. In: IEEE India Annual Conference, INDICON 2004. IEEE, Los Alamitos (2004)
15. Yan, Z., Dong, Y., Niemi, V., Yu, G.: Exploring Trust of Mobile Applications Based on User Behaviors. In: Chen, L., Yung, M. (eds.) INTRUST 2009. LNCS, vol. 6163, pp. 212–226. Springer, Heidelberg (2010)
16. von Reischach, F., Michahelles, F., Guinard, D., Adelmann, R., Fleisch, E., Schmidt, A.: An Evaluation of Product Identification Techniques for Mobile Phones. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5726, pp. 804–816. Springer, Heidelberg (2009)
17. Lee, M., Jeong, H.J.: Mobile Phone Number Management System Using an X-internet Approach. In: Huang, D.-S., Li, K., Irwin, G.W. (eds.) ICIC 2006. LNCS, vol. 4113, pp. 1145–1150. Springer, Heidelberg (2006)
18. Kenteris, M., Gavalas, D., Economou, D.: Mytilene E-guide: a multiplatform mobile application tourist guide exemplar. Springer Science+Business Media, LLC (2010)

Keypoint-Based Enhanced Image Quality Assessment

Ming Zeng[1,2], Jian Liu[1], Youfu Li[2], Qinghao Meng[1],
Ting Yang[1], and Zhengbiao Bai[1]

[1] Institute of Robotics and Autonomous Systems, School of Electrical Engineering and Automation, Tianjin University, Tianjin, China 300072
[2] Department of Manufacturing Engineering and Engineering Management City University of Hong Kong, Hong Kong, China
{zengming,qinghao_meng}@tju.edu.cn,
{morrowindliu,ytyt286,baizhengbiao}@163.com,
meyfli@cityu.edu.hk

Abstract. A simple objective quality assessment for enhanced images is proposed. It requires only scale-space keypoint information. Generally, the distribution of keypoints characterizes structural complexity of local regions. The process of quantitative evaluation involves four steps: The original image is first segmented into two parts (i.e. target region and background). Then, the key points in the original image and the enhanced image are extracted using the strategy of extrema in difference of Gaussian (DoG) scale space. The summed weights (SWs) of keypoints within the target region and background are computed respectively. Finally, the objective assessment result can be achieved by comparing the changes of SWs in two separate regions. The experimental results demonstrate that this proposed algorithm is feasible and effective.

Keywords: image quality assessment, keypoint detection, image enhancement.

1 Introduction

Digital images may be corrupted by some degradation sources, which may arise during acquisition, compression, storage, transmission and reproduction, and have a direct bearing on visual quality. Image enhancement is therefore necessary to improve the visual appearance of the image or to provide a better transform representation for future automated image processing [1]. Although a considerable amount of research has been done on image enhancement, there still does not exist an "off-the-shelf" solution applicable to all types of images. Therefore, the design of reliable quality evaluation algorithm for enhanced images appears to be especially important, it not only can be used to evaluate the performance of various enhancement algorithms, but also can be used to optimize parameter settings of various image enhancement approaches [2].

In general, the existing image quality evaluation approaches can be categorized into two major classes: subjective assessment algorithms, such as mean opinion score (MOS), and objective assessment algorithms [3-6], such as average local variances (ALV), Mean Square Error (MSE) and peak signal-to-noise (PSNR). For subjective

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 420–427, 2011.

assessment (e.g., MOS), quality scores of an image are given by many observers according to some preordained evaluation scale or their own experiences, then weighted-average number is calculated, and the subjective evaluation results are obtained. Although this subjective evaluation method generally has good response to the quality of enhanced image, it is usually too inconvenient, time-consuming and expensive. Objective quality evaluation algorithms, on the other hand, are appealing because they are simple and highly efficient. Note that most of them are known as full-reference, meaning that a complete desirable image is assumed to be known. In many practical applications (e.g. image enhancement), however, the reference image is not available, meaning that the absolute difference of pixels between the original image and the processed image is not highly correlated to the quality difference between them [7]. That is why most of full-reference assessment methods can not be applied to evaluate the quality of enhanced results. For example, the predicted results of above mentioned full-reference quality metrics for the overenhanced image might be better than the results computed for the desirable image.

This paper focuses on no-reference image quality assessment. We develop a new metric, i.e., summed weights of scale-space keypoints, to evaluate the quality of enhanced images. In order to suppress the influences corresponding to noise overenhancement in background, the target region and background region are processed separately and then we will obtain accurate integrated assessment results for enhanced images.

The remainder of this paper is organized as follows. Section 2 will briefly introduce scale-space keypoint detection. In Section 3, we will describe a new keypoint-based quantitative metric for objectively assessing quality of enhanced images. Section 4 will provide assessment results of different testing images. Finally, the conclusions are drawn in section 5.

2 Scake-Space Keypoint Detection

Keypoint detection has been described in details by Lowe in [8]. Here, we briefly review this detection method at an appropriate scale.

In the process of keypoint extraction, the first step is to obtain a set of multi-scale-space images by convolving the original image with variable-scale Gaussian functions:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{1}$$

where $L(x, y, \sigma)$ is the scale-space image, $G(x, y, k\sigma)$ is Gaussian kernel function, σ denotes scale coordinate , and $I(x, y)$ is the input image.

The second step is to detect stable keypoint locations in scale space. Here we use scale-space extrema in the Difference-of-Gaussian (DoG) space, which can be computed from convolution of difference-of-Gaussian kernel of difference scales and the input image:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{2}$$

In order to efficiently find local maxima and minima in the DoG space, each sample point is compared to its eight neighbors in the current image and nine neighbors is the scale above and below. It is selected only if it is larger than all of these neighbors or

smaller than all of them. Moreover, the location and scale of keypoints can be determined accurately by fitting a 3D quadratic function, and at the same time those keypoints that have low contrast, are sensitive to noise are rejected [9-10].

For stability, it is not sufficient to reject keypoints with low contrast since the DoG function will have a strong response along edges. Therefore, those keypoints localized along an edge are also eliminated. Generally, a poorly defined peak in the DoG space will have a large principal curvature across the edge but a small one in the perpendicular direction [11]. The principal curvatures can be computed from a 2×2 Hessian matrix, H:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{3}$$

The eigenvalues of H are proportional to the principal curvatures of D. We can compute the sum of the eigenvalues from the trace of H and their product from the determinant:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \tag{4}$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \tag{5}$$

where α, β are the eigenvalues of H.

To eliminate edge responses, check the ratio between the large magnitude eigenvalue and the smaller one is below threshold, r, we only need to check:

$$\frac{Tr(H)}{Det(H)} < \frac{(r+1)^2}{r} \tag{6}$$

In our proposed method, we use a value of r=10

3 Proposed Scheme

According to the definition of a keypoint in scale space, there must be a structurally distinctive concerned region around the keypoint, which differentiates from the surrounding region. In other words, the distribution of key points characterizes structural complexity of local regions. Here we use this unique feature for evaluating the quality of enhanced images. A new metric, i.e., summed weights (*SWs*) of scale-space keypoints is developed to measure the structural complexity of local regions. In order to suppress the influences corresponding to noise overenhancement in background, the target region and background are processed separately. Generally, the larger *SWs* has the target region along with smaller *SWs* in the background, the better visual appearance exhibits the enhanced image.

The detailed steps are described as follow:

- The original image is first segmented into two parts (i.e., r_t for target region and r_b for background, $R = r_t \cup r_b$). From the many different methods to convert a grayscale image to a binary image we chose a simple, robust method called “Otsu's method”.

- The keypoints (*KPs*) in the original image and the enhanced image are extracted using the strategy of extrema in difference of Gaussian (DoG) scale space. Generally, the value of a scale indicates the importance of its keypoint. Therefore, we set the scales as the weights of *KPs*. For a original image, the summed weights (*SWs*) of keypoints within the target region and background are computed respectively:

$$SW_o(r_t) = \sum_{KP_t \in r_t} w_o(KP_t) \tag{7}$$

$$SW_o(r_b) = \sum_{KP_b \in r_b} w_o(KP_b) \tag{8}$$

where KP_t and KP_b denote the keypoint within the target region and background respectively.

- Similarly, for the enhanced image, the *SWs* of different regions are calculated as follows:

$$SW_e(r_t) = \sum_{KP_t \in r_t} w_e(KP_t) \tag{9}$$

$$SW_e(r_b) = \sum_{KP_b \in r_b} w_e(KP_b) \tag{10}$$

- The changes of SWs in two separate regions are computed as follows:

$$\Delta SW_{eo}(r_t) = SW_e(r_t) - SW_o(r_t) \tag{11}$$

$$\Delta SW_{eo}(r_b) = SW_e(r_b) - SW_o(r_b) \tag{12}$$

- The objective assessment result can be calculated as:

$$Q = \begin{cases} \dfrac{\Delta SW_{eo}(r_t) - \Delta SW_{eo}(r_b)}{SW_o(r_t)}; if\ \Delta SW_{eo}(r_t) \geq \Delta SW_{eo}(r_b) \\ -\dfrac{\Delta SW_{eo}(r_b) - \Delta SW_{eo}(r_t)}{SW_o(r_t)}; otherwise \end{cases} \tag{13}$$

If the value of Q is larger than 0, then it means that the quality of enhanced image is better than that of the original image.

4 Experimental Results

In order to evaluate the effectiveness of the keypoint-based metric, the proposed method has been test on two traditional image enhancement algorithms, i.e., Global Histogram Equalization (GHE) and Contrast-Limited Adaptive Histogram Equalization (CLAHE) [12].

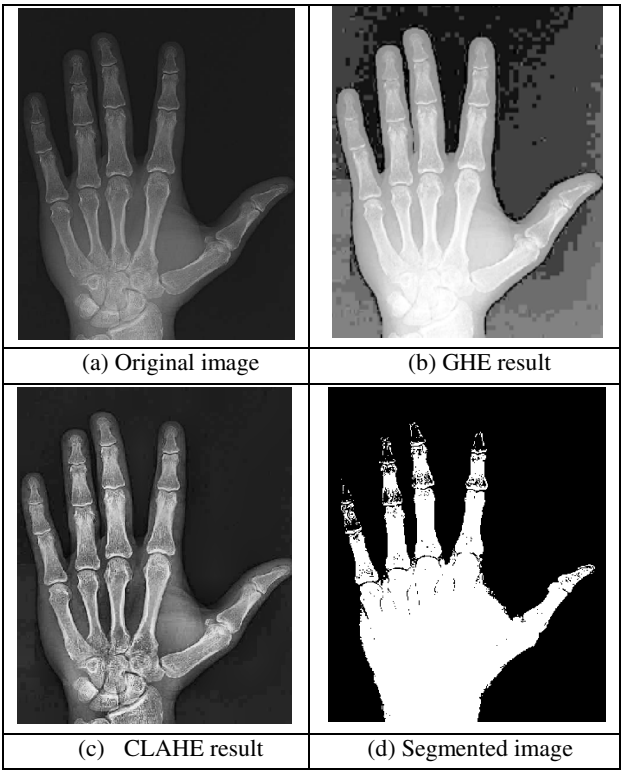


Fig. 1. Comparison of the results for X-ray image of a hand using different enhancement techniques. (a) Original image, (b) Result of GHE, (c) Result of CLAHE, (d) Segmented result for the original image using OTSU.

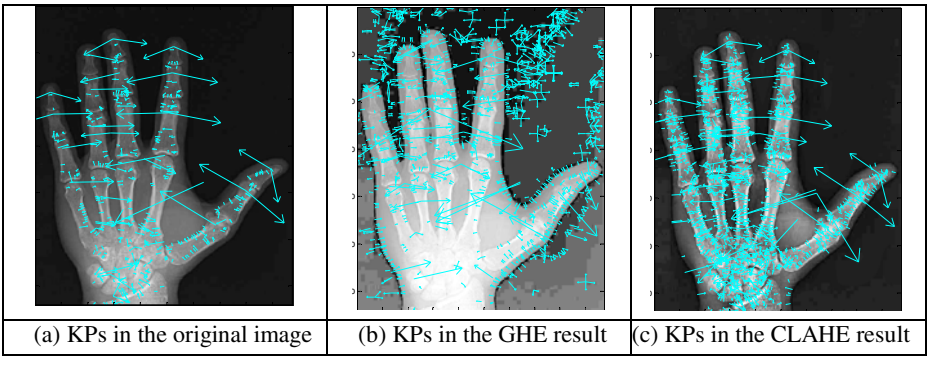
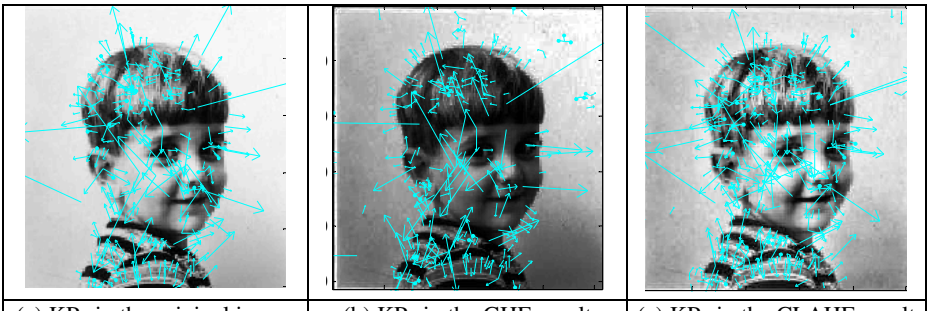


Fig. 2. Distribution of keypoints for Fig1.(a)-(c)

(a) Original image | (b) GHE result | (c) CLAHE result | (d) Segmented image

Fig. 3. Comparison of the results for the little boy image using different enhancement techniques. (a) Original image, (b) Result of GHE, (c) Result of CLAHE, (d) Segmented result for the original image using OTSU.



(a) KPs in the original image | (b) KPs in the GHE result | (c) KPs in the CLAHE result

Fig. 4. Distribution of keypoints for Fig3.(a)-(c)

Fig.1 and Fig.3 illustrate the comparison of the results of treating a X-ray hand image and a little boy image with two enhancement techniques (i.e., GHE and CLAHE). Fig.1(a) and Fig.3(a) are the original test images. Their GHE results shown in Fig.1(b) and Fig.3(b) look faded and background noise have been amplified. Fig.1(c) and Fig.3(c) show the CLAHE results which is better than the GHE results (i.e., Fig.1(b) and Fig.3(b)). The enhanced image (see Fig.1(c) not only exposes more image details, but also has little noise overenhancement.

The distributions of keypoints detected from Fig.1(a)-(c) and Fig.3(a)-(c) are illustrated in Fig.2 and Fig.4 respectively. Fig.1(d) and Fig.3(d) are segmented images in which the white areas are target regions. The objective assessment result can be achieved by analyzing the changes of SWs in two separate regions (i.e. target region and background). The quantitative evaluation results are shown in Table 1 and Table 2. From the Table 1, we can see that the value of Q for GHE is -3.27<0, which means that the quality of the GHE result (i.e., Fig.1(b)) is even worse than that of the original image. Similar results are obtained for the little boy image (see Fig.3(b) and Table 2). In contrast, the value of Q for CLAHE is 1.11>0 meaning that the enhanced image (i.e., Fig.1(c)) has desirable visual appearance compared with the original one (see Fig.1(a)). Table 2 shows that the value of Q for CLAHE result

(Fig.3(c)) is very small ($Q = 0.089$) meaning that CLAHE method does not improve the quality of the original image. It is obvious that these quantitative evaluation results are consistent with subjective assessment results.

Table 1. Quantitative quality assessment results for Fig.1

	GHE	CLAHE
$\Delta SW_{eo}(r_t)$	-258	1276
$(SW_e(r_t) - SW_o(r_t))$	(497- 755)	(2031- 755)
$\Delta SW_{eo}(r_b)$	2211	441
$(SW_e(r_b) - SW_o(r_b))$	(2435- 224)	(665- 224)
Q	**-3.27**	**1.11**

Table 2. Quantitative quality assessment results for Fig.3

	GHE	CLAHE
$\Delta SW_{eo}(r_t)$	-97	98
$(SW_e(r_t) - SW_o(r_t))$	(275- 372)	(470-372)
$\Delta SW_{eo}(r_b)$	25	65
$(SW_e(r_b) - SW_o(r_b))$	(388-363)	(428-363)
Q	**-0.33**	**0.089**

5 Conclusion

Objectively assessing the quality of enhanced images is a challenge problem since the reference image is not available. From various experiments, we have found that the distribution of keypoints is a useful cue to measure the structural complexity of local regions. Based on this assumption, we develop a keypoint-based measurement technique to evaluate the quality of enhanced images. In contrast to existing full-reference quality assessment methods in which the process depends on the entire image, our method introduces a double-region strategy to compare the relative enhancement appearance of two separate regions (i.e., target region and background) because many image enhancement algorithms suffer from noise over-enhancement in background. Test results on two traditional image enhancement algorithms (i.e., GHE and CLAHE) show that this model's predictions agree well with outputs of our visual perception.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant Nos.60802051, 60875053, in part by Tianjin Natural Science Foundation under Grant No. 09JCYBJC02100, in part by Youth Foundation of Tianjin University, and in part by China '863' High-Tech Program

under Grant No. 2007AA04Z219. M. Zeng would like to thank the support of the Research Grants Council of Hong Kong (project No. CityU 117507).

References

1. Agaian, S.S., Silver, B., Panetta, K.A.: Transform coefficient histogrambased image enhancement algorithms using contrast entropy. IEEE Transactions on Image Processing 16, 16, 741–758 (2007)
2. Eskiciogln, A.M.: Image Quality Measure and their performance. IEEE Transactions on Communications 43, 2959–2965 (1995)
3. Eckert, M.P., Bradley, A.P.: Perceptual quality metrics applied to still image compression. Signal Processing 70, 177–200 (1998)
4. Bouzerdoum, A., Havstad, A.: Image quality assessment using a neural network approach. In: Proceedings of the Forth IEEE International Symposium on Signal Processing and Information, Rome, Italy, pp. 330–333 (2004)
5. Mannos, J.l., Sakrison, D.J.: The effects of a visual fidelity criterion on the encoding of images. IEEE Transaction on Information Theory 4, 525–536 (1974)
6. Russo, F.: An image enhancement technique combining sharpening and noise reduction. IEEE Transactions on Instrumentation and Measurement 51, 824–828 (2002)
7. Wang, Z., Bovik, A.C., Sheikh, H.R.: Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing 13 (2004)
8. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Key points. International Journal of Computer Vision 60, 91–110 (2004)
9. Lowe, D.G.: Object recognition m local scale-invariant features. In: Seventh International Conference on Computer Vision, Kerkyra, Greece, pp. 1150–1157 (1999)
10. Wang, Z.Z., Jia, K.B., Liu, P.Y.: A Novel Image Retrieval Algorithm Based on ROI by Using SIFT Feature Matching. In: Proceedings of International Conference on MultiMedia and Information Technology, pp. 338–341 (December 2008)
11. Brown, M., Lowe, D.G.: Invariant Features from interest point groups. In: Proceedings of the 13th British Machine Vision Conference, pp. 656–665 (2002)
12. Pizer, S.M., Amburn, E.P., Austin, J.D.: Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing 39, 355–368 (1978)

Finding Regions of Interest Based on Scale-Space Keypoint Detection

Ming Zeng[1,2], Ting Yang[1], Youfu Li[2], Qinghao Meng[1], Jian Liu[1], and Tiemao Han[1]

[1] Institute of Robotics and Autonomous Systems, School of Electrical Engineering and Automation, Tianjin University, Tianjin, China 300072
[2] Department of Manufacturing Engineering and Engineering Management City University of Hong Kong, Hong Kong, China
{zengming,qinghao_meng}@tju.edu.cn,
{ytyt286,morrowindliu}@163.com, meyfli@cityu.edu.hk,
tmhan@mail.tju.edu.cn

Abstract. One of the major challenges for modeling visual attention mechanisms is to extract visual cues for automatic detection perceptually important regions in a scene. Here, we propose a simple model for detecting regions of interest (ROI) inspired from keypoint analysis. We adopted the idea that the appearance of an interest region can be well characterized by the distribution of its local features (e.g., keypoints). ROI detection involves five steps: the input image is first decomposed into a set of low-level vision feature maps (e.g., intensity map, and two double-opponent color maps). Extrema in difference of Gaussian (DoG) scale space are then calculated for detecting the keypoints within each feature maps. The location and scale information of keypoints are integrated to create three conspicuity maps. These conspicuity maps are normalized and summed into an overall saliency map. Finally, a "small" number of salient locations are successively selected using a dynamical neural network. Experimental results show that the proposed model outperforms the Itti's model, a state-of-the-art competitive approach.

Keywords: regions of interest, scale-space keypoint detection, saliency map.

1 Introduction

Selective visual attention is a critical mechanism for scene analysis, which enables one to allocate the limited perceptual resources in an efficient way. By means of attention, it is possible to break down the analysis of a complex scene to the analysis of its parts through a succession of fixations and saccades. Those parts of the image that cause eye fixations and capture primary attention are called regions of interest (ROI). Detecting ROI in a complex image is challenging but useful in many image processing applications, such as intelligent image compression [1], high-speed object recognition [2], content-based image retrieval [3], medical image analysis [4,5] and active vision.

In most cases, a likely ROI candidate generally has visually and structurally distinctive features compared with its neighboring region. Therefore, some structural

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 428–435, 2011.

characteristics, such as contrast, color, region size and shape, distribution of contours or texture pattern, can be used to detect the ROIs. Most of the developed algorithms are based on this assumption. To mimic the perceptual grouping mechanism in the human visual system, Lau and Levine select a number of biologically motivated features for representing the visual properties of a region. These features were color (L*a*b*), texture (Gabor), and position. [6] A selective visual attention model proposed by Aziz and Mertsching computes five properties for each region: color, size, symmetry, orientation and eccentricity. The rarity or contrast of these values determines the feature saliency values, which are then combined into a single salience map.[7] Inspired by the behavior and neuronal architecture of the early primate visual system, Itti and Koch used the strategy of multiscale centre-surround contrast in terms of intensity, color and orientation to find several salient locations in a scene.[8] Kootstra and Schomaker used local symmetry operators to measure salence, which they found to be a better match to experimental human data than Itti and Koch's contrast-based saliency measure.[9] Kadir and Brady used a local measure of entropy to detect salient regions in an image.[10] The method proposed by Caron et al. used power laws (e.g., Zipf law and inverse Zipf law) to measure the frequency of appearance of the patterns contained in images, which allows to characterize the structural complexity of image textures. This complexity measure indicates a perceptually salient region in the image.[11] A related recent approach by Rosin used only edge information to detect salient regions. To avoid requiring any parameters to be set they use the distance transform (DT) to propagate the edge information. However, since the standard DT is defined for binary images rather than grey level images they apply threshold decomposition. That is, all the possible threshold images (e.g., 66 images) are generated from the grey level image, each is processed separately, and then the set of DT maps is summed to form the salience map. It is obvious that the computational cost is very high. In particular, since it is edge based then it will fail if many strong edges exist in the image that do not belong to salient objects, e.g. from strongly textured backgrounds.[12]

Although a considerable amount of research has been done on ROI detection, there still does not exist an “off-the-shelf” solution applicable to all types of images. A major issue has been the lack of a good measure of importance of a local region. From many experiments, we found that the appearance of an interest region can be well characterized by the distribution of keypoints. In this paper, we propose a simple keypoint-based ROI detection method. It is extremely simple to implement, it requires only the position and scale information of scale-space keypoints.

The remainder of this paper is organized as follows: Section 2 introduces the basic principles of our proposed model including feature extraction, keypoint detection, saliency map and ROI detection. Experimental results are reported in Section 3. Finally, conclusions are drawn in section 4.

2 Model Description

In general, scale-space keypoints are highly distinctive, relatively easy to extract, allow for correct object tracking with low probability of mismatch and are easy to match against a large database of local features. With these advantages, the strategy of

keypoint detection has been applied to object tracking, object recognition, and some biomedical applications. Here we use this unique feature for finding a "small" number of ROI in a scene.

ROI detection follows a few typical steps:

2.1 Extracting the Primary Feature Maps

Suppose that the input image is a color image. The input image is composed of three channels defined as r, g, and b, which represent red, green, and blue values, respectively. Intensity image M_I is obtained by averaging the red, green, and blue color channels as $M_I = (r+g+b)/3$.

Color is obviously a potentially very useful property. Previous studies suggest color opponency can help support observations. From the r, g, and b channels we extract the red-green (RG) and blue-yellow (BY) maps. To extract two opponent color maps, we use the definition proposed in Walther and Koch (2006) [13] which gives better results than those in Itti et al. (1998) [8]. The RG and the BY maps are defined as follows:

$$M_{RG} = \frac{r-g}{\max(r,g,b)} \tag{1}$$

and,

$$M_{BY} = \frac{b-\min(r,g)}{\max(r,g,b)} \tag{2}$$

Moreover, in order to avoid the hue instability when the intensity level is low, M_{RG} and M_{BY} are set to zero when $\max(r, g, b) < 0.1$ (i.e., all the red, green and blue components are less than a tenth of the dynamic range).

2.2 Scale-Space Keypoint Detection

Keypoint detection has been described in details by Lowe in [14]. Here, we briefly review this detection method at an appropriate scale.

In the process of keypoint extraction, the first step is to obtain a set of multi-scale-space images by convolving the original image with variable-scale Gaussian functions:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{3}$$

where $L(x, y, \sigma)$ is the scale-space image, $G(x, y, k\sigma)$ is Gaussian kernel function, σ denotes scale coordinate , and $I(x, y)$ is the input image.

The second step is to detect stable keypoint locations in scale space. Here we use scale-space extrema in the Difference-of-Gaussian (DoG) space, which can be computed from convolution of difference-of-Gaussian kernel of difference scales and the input image:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{4}$$

In order to efficiently find local maxima and minima in the DoG space, each sample point is compared to its eight neighbors in the current image and nine neighbors is the scale above and below. It is selected only if it is larger than all of these neighbors or smaller than all of them. Moreover, the location and scale of keypoints can be determined accurately by fitting a 3D quadratic function, and at the same time those keypoints that have low contrast, are sensitive to noise are rejected [14].

For stability, it is not sufficient to reject keypoints with low contrast since the DoG function will have a strong response along edges. Therefore, those keypoints localized along an edge are also eliminated. Generally, a poorly defined peak in the DoG space will have a large principal curvature across the edge but a small one in the perpendicular direction. The principal curvatures can be computed from a 2×2 Hessian matrix, H:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{5}$$

The eigenvalues of H are proportional to the principal curvatures of D. We can compute the sum of the eigenvalues from the trace of H and their product from the determinant:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \tag{6}$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \tag{7}$$

where α, β are the eigenvalues of H.

To eliminate edge responses, check the ratio between the large magnitude eigenvalue and the smaller one is below threshold, r, we only need to check:

$$\frac{Tr(H)}{Det(H)} < \frac{(r+1)^2}{r} \tag{8}$$

In our proposed method, we use a value of r=10.

2.3 Saliency Map and ROI Detection

We carry out the above scale-space analysis on the color (i.e., M_{RG} and M_{BY}) and intensity (i.e., M_I) channels of an image. This results in the identification of interesting points and the associated scales. According to the definition of a keypoint in scale space, there must be a structurally distinctive concerned region around the keypoint, which differentiates from the surrounding region. In other words, the distribution of key points characterizes structural complexity of local regions.

The area of the concerned region is simply a square window that is centered at the interesting point and whose size is determined by the scale. This implies that the windows can overlap. In order to label the importance of this concerned region, a weighting factor for each pixel over within it is set to 1, and the rest of the region is set to 0. Each keypoint is processed separately, and then the results are summed to produce a single conspicuity map for each feature. Next, three conspicuity maps are

normalized and linearly combined into an overall saliency map (SM). Finally, the SM feeds into a biologically plausible 2D "winner-take-all" (WTA) neural network [8,15] to select a small set of salient locations. The overall flowchart of ROI detection is illustrated in Fig.1.

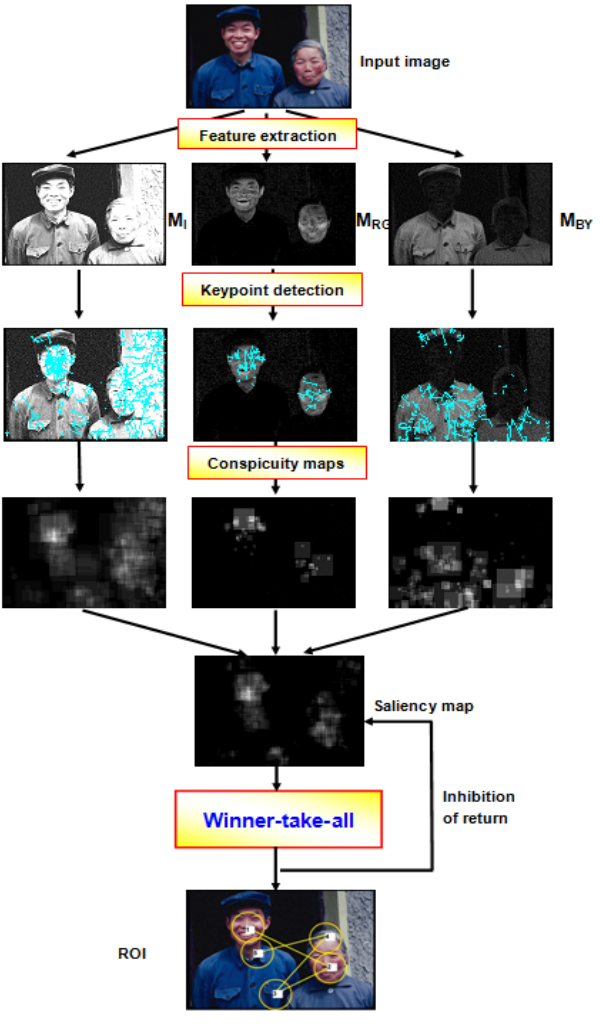
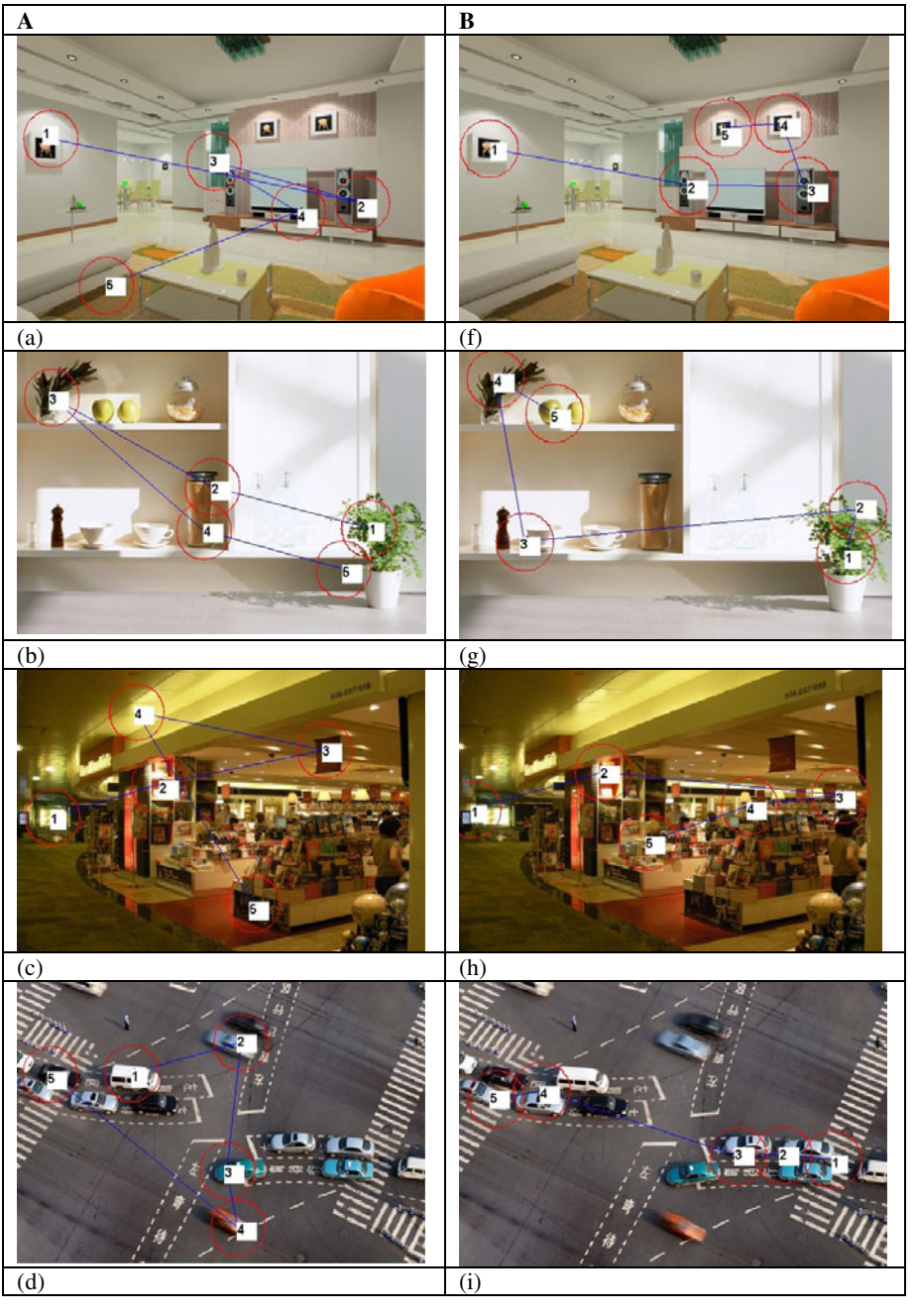
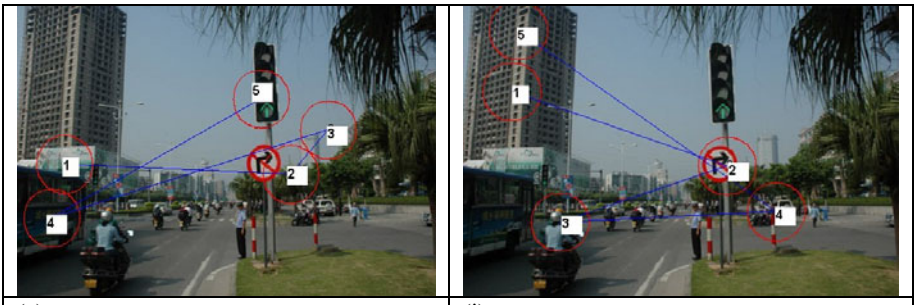


Fig. 1. Flowchart of the keypoint-based ROI detection model

3 Experimental Results

To verify the effectiveness of our proposed model, we tested our ROI detection algorithm on many indoor and outdoor natural images. Figure 2 shows several examples of the detected ROI by our algorithm along with results from Itti and


A
B
(a)
(f)
(b)
(g)
(c)
(h)
(d)
(i)



(e) (j)

Fig. 2. Comparison of different models for several natural images. Column A illustrates gaze-shift results of our keypoint-based model (i.e., (a)-(e)). Column B shows gaze-shift results of Itti's model (i.e., (f)-(j)).

Koch's model for comparison. In these figures, five attended areas are selected in order of decreasing saliency in which the red circle represents the border of the ROI and blue lines illustrate the dynamical shifts of the ROI.

The results of Itti's algorithm are shown in Fig.2a-Fig.2e. The results of our proposed algorithm are shown in Fig.2f-Fig.2j. From the comparison results, some general comments can be made. First, using Itti's model, some important regions are missed and localization is poor, e.g., as seen in Fig.2a two ornamental paintings on the wall which are regarded as significant objects according to our visual perception are not included in the top five salient locations. In Fig.2e, the Itti's model excludes some conspicuous locations such as the tall building, passersby and vehicles on the street. Second, the dynamical shifts of the focus of attention (FOA) predicted by Itti's model exhibit a lack of order or regular arrangement, as shown in Fig.2b-Fig.2d. In all test images the proposed method has successfully located all the main objects, and rejected the majority of the background. In addition, our model produced more accurate estimation of the gaze shifts than the Itti's model.

4 Conclusion

A simple model for ROI extraction using the keypoint analysis is proposed in this paper. This automatic extraction of ROI is able to be applied to any objects. The proposed model has promising signification in some related research fields such as image compression, video surveillance and human-computer interaction. There are several possible remaining issues for further investigation of ROI detection. For future work, we are interested in extracting the integral interesting objects automatically, i.e., object-based ROI detection. We believe that the model considering about the edge of the objects could perform better. In addition, we plan to investigate the problem of optimal gaze shifts which agree well with human subjective perception.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grant Nos.60802051, 60875053, in part by Tianjin Natural Science Foundation under Grant No. 09JCYBJC02100, in part by Youth

Foundation of Tianjin University, and in part by China '863' High-Tech Program under Grant No. 2007AA04Z219. M. Zeng would like to thank the support of the Research Grants Council of Hong Kong (project No. CityU 117507).

References

1. Han, S.Y., Vasconcelos, N.: Image Compression using Object-based Regions of Interest. In: Proceedings of IEEE International Conference on Image Processing, pp. 3097–3100 (2006)
2. Vazquez-Martin, R., Marfil, R., Nunez, P., et al.: A Novel Approach for Salient Image Regions Detection and Description. Pattern Recognition Letters, 1464–1476 (2009)
3. Vu, K., Hua, K.A., Tavanapong, W.: Image Retrieval Based on Regions of Interest. IEEE Transactions on Knowledge and Data Engineering 15(4), 1045–1049 (2003)
4. Hoge, W.S., Miller, E.L., Lev-Ari, H.: An Efficient Region of Interest Acquisition Method for Dynamic Magnetic Resonance Imaging. IEEE Transactions on Image Processing 10(7), 1118–1128 (2001)
5. Jahanbin, R., Sampat, M.P., Muralidhar, G.S., et al.: Automated Region of Interest Detection of Spiculated Masses on Digital Mammograms. In: Proceedings of IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 129–132 (2008)
6. Lau, H.F., Levine, M.D.: Finding a Small Number of Regions in An Image using Low-level Features. Pattern Recognition 35, 2323–2339 (2002)
7. Aziz, M.Z., Mertsching, B.: Fast and Robust Generation of Feature Maps for Region-based Visual Attention. IEEE Transactions on Image Processing 17(5), 633–644 (2008)
8. Itti, L., Koch, C., Niebur, E.: A Model of Saliency-based Visual Attention for Rapid Scene Analysis. IEEE Transactions on Pattern Analysis Machine Intelligence 20(11), 1254–1259 (1998)
9. Kootstra, G., Nederveen, A., Boer, B.: Paying Attention to Symmetry. In: Proceedings of the British Machine Vision Conference, pp. 1115–1125 (2008)
10. Kadir, T., Brady, M.: Saliency Scale and Image Description. International Journal of Computer Vision 45(2), 83–105 (2001)
11. Caron, Y., Makris, P., Vincent, N.: Use of Power Law Models in Detecting Region of Interest. Pattern Recognition 40, 2521–2529 (2007)
12. Rosin, P.L.: A Simple Method for Detecting Salient Regions. Pattern Recognition 42, 2363–2371 (2009)
13. Walther, D., Koch, C.: Modeling Attention to Salient Proto-objects. Neural Networks 19, 1395–1407 (2006)
14. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110 (2004)
15. Tsotsos, J.K., Culhane, S.M., Wai, W.Y.K., et al.: Modelling Visual Attention via Selective Tuning. Artificial Intelligence 78(1-2), 507–545 (1995)

SVM Classifier of Stored-Grain Insects Based on Grid Search

Zhang Hongtao[1], Yin Shuping[2], and Hu Yuxia[2]

[1] Institute of Electric power, North China University of Water Conservancy and Electric Power, Zhengzhou 450011, China
[2] College of Electric Engineering, Zhengzhou University, Zhengzhou 450001, China
zht1977@ncwu.edu.cn

Abstract. The detection of the stored-grain insects based on image recognition technology is high accuracy, cost-effective, high efficiency, no pollution and less labor. The classification of the stored-grain insects was multi-feature and multi-compound degree of various insects. The classifier design was critical for the detection system of the stored-grain insects. The optimal parameters C and g should be identified while using RBF kernels in the SVM classifier. The goal was to get the best cross-validation training model and improve the classification accuracy of the classifier. The grid search consisting of rough selection and fine selection was proposed to optimize parameters C and g by the recognition ratio of the training model. The optimal parameters were 30574 and 0.5743 after training, respectively. The fifteen species of the stored-grain insects that spoiled seriously in granary were automatically recognized by SVM classifier, and the correct identification ratio was over 94.67%. The experiment showed that it was practical and feasible.

Keywords: Stored-grain insects, image recognition, support vector machine, grid search, feature extraction.

1 Introduction

Overall post-harvest losses of grains of about 10–15% are fairly common in the world. The stored-grains of at least 5% are spoiled by the stored-grain insects every year[1]. The stored-grain insects cause quantity and quality losses and lower grain values. Insects not only consume grain but also contaminate it with their metabolic byproducts and body parts. Heavily infested stored grains are unfit for seed purposes and its products are unsuited for human consumption[2]. The efficient detection for the stored-grain insects is very critical for predicting the development of insect populations and for making proper management decisions to control them below the economical threshold. Moreover, the reliable detection can reduce the stored-grain losses and the pollution to the stored grains and the environment.

The accurate identification of the stored-grain insects is the kernel of the real-time detection and the forecast and prediction. Many different methods are used to recognize the stored-grain insects, for instance, the naked-eye recognition, the acoustic recognition and the image recognition. The image recognition has been hot in

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 436–441, 2011.

research for recent ten years, and it has been shown to have great potential for detecting and identifying various stored-grain insects such as high accuracy, low price, high efficiency and no pollution.

Zayas and Flinn identified insects and body parts of *Rhyzopertha dominica* in bulk wheat samples with subimages of 8×8 pixels based on digital imaging techniques. Recognition success of *R. Dominica* adults, some grass seeds, and other non-wheat components were higher than 90%[3]. Ridgway developed a machine vision system for the detection of adult beetles, rodent droppings and ergot in harvested wheat. Detection rates for artificially prepared samples ranged from 87% for ergot to 100% for adult *Oryzaephilus surinamensis*. The speed of the integrated algorithm package for detection of all three contaminant types approaches that required to scan 3 kg grain in just 3 min[4]. However, the real-time detection system of the stored-grain insects based on machine vision need higher automation, and the species of the stored-grain insects to be recognized should be increased further.

The classification of the stored-grain insects is multi-feature and multi-compound degree of various insects. The classifier design is the key to the automatic detection system of the stored-grain insects. Support Vector Machine (SVM) is the new generation machine learning algorithm which develops on the basis of the statistics learning theory. It has better classified ability and generalization ability to solve pattern recognition problems that are few-sample, non-linear and high dimensions. It has also been proved to be very successful in many other classification applications such as plant disease-spots, mechanical failures and smell analysis[5]. The grid search is proposed to optimize parameters C and g judged by the cross-validation training model accuracy in the SVM classifier. The fifteen species of the stored-grain insects are automatically recognized by the SVM classifier. The experiment showed that it was practical and feasible.

2 SVM Classification Theory

SVM is a state-of-the-art pattern recognition technique whose principles stem from the minimal empirical error in the statistical learning theory[6]. It was first introduced as large margin classifiers. For a separable training set, the margin is defined as the minimum distance between the points of two classes, measured perpendicularly to the separating hyperplane. Maximizing this margin is a way for a learning algorithm to control the ability and the complexity of the machine, and to select the optimal separable hyperplane among all the hyperplanes that separate the training set into two classes. Therefore, the SVM classification can be formulated as follows: The training samples are well classified by maximizing the margin[7].

Given the training sample set (x_i, y_i), where $x_i \in R^d$ is a feature vector, $i = 1, \cdots, n$, for the binary classification, $y_i \in \{-1,1\}$ is the corresponding target. In fact, the classification may be converted to the following quadratic programming problem[8]:

$$\text{minimize} \quad \frac{1}{2}(w,w) + C\sum_{i=1}^{n}\xi_i \tag{1}$$

$$\text{subject to} \quad y_i[(w, x_i)+b] \geq 1-\xi_i, \ \xi_i > 0 \tag{2}$$

where (,) is the dot product between two vectors, a set of positive slack variables ξ_i are introduced when the training samples are not linearly separable. C is a positive regularization factor, it controls the magnitude of the penalty associated with training samples that lie on the wrong side of the hyperplane. w and b are the weight vector and the classification threshold in the decision function $f(x)=(w,x)+b$, respectively.

The problem can be turned into maximizing the following dual problem:

$$\text{Maximize} \quad \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j K(x_i, x_j) \tag{3}$$

$$\text{subject to} \quad 0 \leq \alpha_i \leq C, \ \sum_{i=1}^{n} y_i\alpha_i = 0, i = 1, \cdots, n \tag{4}$$

where α_i is the lagrange multiplier corresponding to every training sample, and the magnitude of α_i determined by the parameter C. It is a quadratic function optimization problem under the inequality constraint, the unique solution exists and the no-zero α_i in the solution corresponds to the support vector. The optimization classified function is obtained after solving the above question, it also is the optimal decision function in fact.

$$f(x) = \text{sgn}[\sum_{r=1}^{s}\alpha_i y_i K(x_r, x) + b] \tag{5}$$

Where sgn( ) is the sign function, x_r is the rth support vector. Combining the equation $w = \sum_{i=1}^{n}\alpha_i y_i x_i$ with $y_i[(w \cdot x_i)+b]-1=0$, the threshold b can be solved by importing any support vector. The non-linear translation defined by the dot product function $K(x_r, x)$ maps the input data space into a higher-dimensional Hilbert space and attempts to seek the optimization linear separable hyperplane in the higher-dimensional space. Frequently used kernel functions are the polynomial kernel, the spline function kernel and the radial basis function kernel. The radial basis function kernel $K(x_r, x) = \exp(-g \| x_r - x \|^2)$ is used widely, where g>0.

The binary classification is the most basic classification strategy of SVM. And the binary approach has the clear geometry meaning, the algorithm is quite simple and perfect. The binary classification of SVM need be extended to the multi-classification in practice. There are various algorithms such as one-against-one and one-against-all.

3 SVM Parameters Optimization Based on Grid Search

The Libsvm Software Kit developed by Lin Zhiren can be used to resolve the SVM classification of 15 species of stored-grain insects (7 features) under the Matlab environment[9]. The penalty factor C and g in the radial basis function kernel need be

identified in the SVM classifier. The ranges of C and g are $(2^{-5},2^{15})$ and $(2^{-16},2^{4})$, respectively. Therefore, it is difficult to seek for the optimization in the wide scopes.

The accuracy with which SVM may classify a data set is dependent on the magnitude of the parameters C and g. With a large value of C and g, there is a tendency to over-fit the training data for SVM, yielding a classifier that may generalize poorly. The generalization ability of the classifier is of fundamental importance and the over-fitting is undesirable. Consequently, the values of C and g must be determined carefully.

The grid search is a commonly used parameter optimization method, and it includes two steps: the rough selection and the fine selection. Firstly, the rough selection result gives a approximate fixed parameter range. Secondly, the fine search is implemented with a smaller step in the range determinted by the rough selection result. To reduce the subjectivity in setting the SVM parameters, the cross-validation is adopted for their determination. In fact, C and g are optimal when the cross-validation training model accuracy is the highest.

In the parameters optimization of the SVM classifier for the stored-grain insects, the criterion is the cross-validation training model accuracy $J(c_i,g_i)$. Where $J(c_i,g_i)$= *svmtrain*(*training_label_vector, training_instance _matrix*,'-c c_i–g g_i -v 5'), *training _label_vector* is the label vector of the training samples, *training_instance _matrix* is the feature matrix of the training samples. The parameters c_i and g_i are C and g in the ith loop, respectively. The parameter v is the folds in the cross-validation.

The v-fold cross-validation means that the training samples are randomly splitted into v parts. The v-1 parts are the training data set, and the remaining part is the validation data set with which validates the classifier accuracy. The v loops are implemented until the v parts data are selected entirely. *Svmtrain* is the mean of the v cross-validation accuracies.

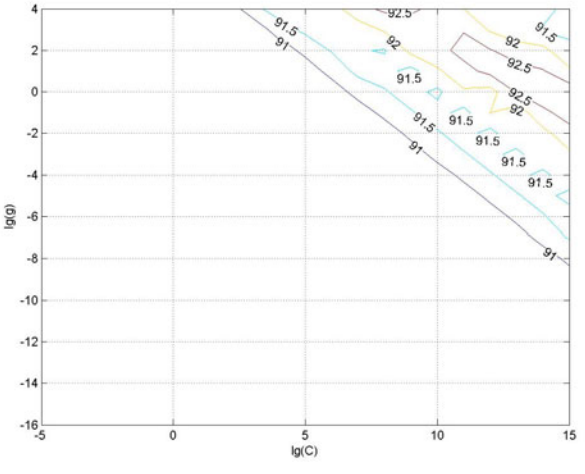


Fig. 1. Rough-selection result of grid search for parameters C and g

The SVM classifier is trained by 1500 samples of the stored-grain insects (100 every species), and the loop step is 2^1.The rough selection result is shown in Figure 1, the coordinates are the logarithm of C and g, respectively. The contour is the line linked between the next-dot with equal cross-validation accuracy displayed on the line. The cross-validation accuracy reaches to the highest at 92.5%. According to the calculation complexity of the fine selection and the optimization quality, the optimal results of the rough selection are $C \in (2^{7.5},2^{15})$ and $g \in (2^{-2},2^{4})$. The ranges of the fine selection are the rough selection results, and the loop step is $2^{0.1}$. Figure 2 shows the fine selection result, and the *svmtrain* reaches to the highest at 93% containing only two parts. Thus, the optimization parameters are $C=2^{14.9}=30574$ and $g=2^{-0.8}=0.5743$.

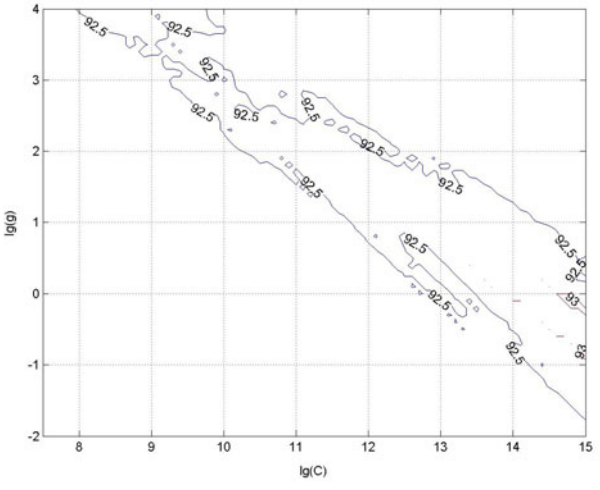


Fig. 2. Fine-selection results of grid search for parameters C and g

4 SVM Classification

The recognition of fifteen species of stored-grain insects belongs to the multi-classification. The Libsvm Software Kit adopts one-against-one classification algotithm, and it turns the multi-classification into the classical binary classification. Combined with any two species, the classification of fifteen species of stored-grain insects totally obtains 105 binary-classifiers. The poll strategy is utilized to analyze the classification results. The stored-grain insect to be recognized belongs to the species that the poll are the highest.

The SVM classifier has 750 support vectors after the training of 1500 samples. The SVM classifier validates 750 samples (50 every species). The number of the support vectors of the classifier is 343. More support vectors imply the multi-compound degree among the species, and it is difficult to classify them rightly.

The 40 of 750 samples are wrongly classified, so the recognition accuracy is 94.67%. This is because that the base learning theory of the SVM classifier based on the statistics learning theory is to adapt with the limited training samples. Consequently, it has the better classified ability in solving the few samples and non-linear recognition questions.

5 Conclusions

The fifteen species of the stored-grain insects that spoil seriously in granary are classified by the SVM classifier with seven features. Having been trained and tested with some samples, the classifier accuracy is 94.67%. It is confirmed that the SVM classifier can achieve anticipated results to the recognition of fifteen species of the stored-grain insects. The intelligence detection system realized the real-time counting and the automatic classification of the stored-grain insects. How to identify the insect in grain kernels automatically needs to be researched further.

Acknowledgement

Financial support was provided by the National Natural Science Foundation of China (No. 30871449) and the Natural Science Research Program of the Education Department of Henan Province (No. 2011B210028).

References

1. Neethirajan, S., Karunakaran, C., Jayas, D.S.: Detection Techniques for Stored-product Insects in Grain. Food Control 18(2), 157–162 (2007)
2. Singh, C.B., Jayas, D.S., Paliwal, J.: Detection of Insect-damaged Wheat Kernels Using Near-infrared Hyperspectral Imaging. Journal of Stored Products Research 45(3), 151–158 (2009)
3. Ridgway, C., Davies, E.R., Chambers, J., Mason, D.R., Bateman, M.: Rapid machine vision method for the detection of insects and other particulate bio-contaminants of bulk grain in transit. Biosystems Engineering 83(1), 21–30 (2002)
4. Zayas, I.Y., Flinn, P.W.: Detection of Insects in Bulk Wheat Samples with Machine Vision. Transactions of the ASAE 41(3), 883–888 (1998)
5. Fung, Y.H., Chan, Y.H.: A Simulated Annealing Restoration Algorithm for Restoring Halftoned Color-quantized Images. Signal Processing: Image Communication 21(4), 280–292 (2006)
6. Buciu, I., Kotropoulos, C., Pitas, I.: Demonstrating the Stability of Support Vector Machines for Classification. Signal Processing 86(9), 2364–2380 (2006)
7. Chen, X.W., Zeng, X.Y., Alphen, D.V.: Multi-class Feature Selection for Texture Classification. Pattern Recognition Letters 27(14), 1685–1691 (2006)
8. Chen, G.Y., Xie, W.F.: Pattern recognition with SVM and Dual-tree Complex Wavelets. Image and Vision Computing 25(6), 960–966 (2007)
9. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines (2002), `http://www.csie.ntu.edu.tw/~cjlin/libsvm`

Pattern Recognition of Handwritten Text Based on Bayes Algorithm

Jianming Cui [1,2]

[1] School of Information Engineering, Chang'an University, Xi'an, China
cjianming@gmail.com
[2] Shaanxi Road Traffic Detection and Equipment Eng. Research Center, Xi'an, China

Abstract. In this paper, have been described using pattern recognition, is also introduced for use Bayes algorithm in pattern recognition. Pattern recognition is the machine identification, computer identification or machine Automatic Identification, aims is to automatically identify things in the machine. After the introduction of the Bayes algorithm, pattern recognition accuracy is improved, and the use of a template approach, so the speed has also been taking into account.

Keywords: Pattern recognition, Bayes Algorithm, Recognition of Handwritten Text.

1 Instruction

Extensive use of pattern recognition[1], such as handwritten digit recognition, intelligent traffic control signal recognition, text recognition, and speech recognition and so on. The purpose of interdisciplinary research is to make the machine do what humans can do; with human beings have analysis, description all kinds of things and phenomena and some ability to judge. Pattern recognition is an intuitive, ubiquitous, humans and animals with pattern recognition are extremely common, but it is very difficult for the computer. The machine can identify, classify things, it need to study the method. Pattern recognition can be summarized into two types, one is intuitive image, such as pictures, photos, graphics, character design, etc. show Figure1,what is it?; The other is not only perceived image of the waveform data or signals, such as language, sound, ECG, seismic waves and so on. Character recognition processing information can be divided into two categories:the one is the text type of information, mainly dealing with the text written or printed text of countries and nations ; the other is the data type of information, constitutes the various numbers and statistics data by mainly Arabic numerals and a small amount of special symbols .

Fig. 1. Images to be recognition

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 442–447, 2011.

Handwritten numeral recognition is a character recognition processing, object of study is how use of computer automated to identify the human figures Arabic numerals by the hand. In the paper, research that following the introduction of Bayesian formula[2,3].

2 Bayes Algorithm

If known the total M class of objects, and a variety of the statistical distribution in d-dimensional feature space. Specifically, the known each class w_i= 1,2,…, M the prior probability $P(w_i)$ and various types of conditional probability density function $P(w_i/X)$, for tested samples, by Bayes formula can be calculated the probability of the sample in their respective categories, called the posterior probability,

Posterior probability as a basis for identifying the ownership of the object to see X has the largest probability in which class, and put the X attributable to that most likely class.

Bayes formula can be expressed as,

$$P(w_i/X) = \frac{P(X/w_i)}{\sum_{j=1}^{n} p(X/w_i)/p(w_i)} \tag{1}$$

Where: $P(w_i/X)$ is called the posterior probability.

Analog state is a random variable, and the probability that a state can be estimated. Bayes formula reflects the relationship between the prior probability, the class probability density function and the posterior probability.

2.1 Prior Probability

For a time to the possibility of M, the prior probability $P(w_i)$ does not consider any other conditions. If the total number is n, where the number of categories 1 is n1, the number of class 2 is n2, then:

$$P(w_1) = \frac{n_1}{n}, \quad P(w_2) = \frac{n_2}{n} \tag{2}$$

2.2 Class Conditional Probability Density Function $P(X/w_i)$

Class conditional probability density function P(*X*/w_i) refers to appear the probability density of the feature X in the feature space of one class, namely, is how its properties *X* of w_i category distributed.

Many of the problems in engineering, statistical data often satisfy normal distribution. Normal distribution is simple, convenient, less parameters, and is a suitable mathematical model. If using normal density function as the function form of the class conditional probability density, within the parameters of the function are unknown, such as such as the expected value and variance, Then this question becomes that how to use a large number of samples to estimate these parameters, As

long as these parameters are estimated, the class conditional probability density function P(X/w_i) also determined. The univariate normal probability density function is:

$$P(X) = \frac{1}{2\pi\sigma} exp\left[-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2\right] \tag{3}$$

Where,

μ ——is the mathematical expectation;

σ ——for the variance.

Multi-dimensional normal probability density function:

$$P(X) = \frac{1}{(2\pi)^{\frac{N}{2}}|S|^{\frac{1}{2}}} exp\left[-\frac{1}{2}(x-\bar{\mu})^T S^{-1}(x-\bar{\mu})\right] \tag{4}$$

Where,

S—— is the N-dimensional covariance matrix;

S-1——is the inverse matrix of S;

|S|——is the determinant of S;

$\bar{\mu}$——is the N-dimensional mean vector.

In most cases, the class conditional probability density functions can use multi-dimensional variable density function of normal function to simulate, namely,

$$P(X/w_i) = ln\frac{1}{(2\pi)^{\frac{N}{2}}|S_i|^{\frac{1}{2}} exp\left[-\frac{1}{2}(X-\bar{\mu}_i)^T S_i^{-1}(X-\bar{\mu}_i)\right]} \tag{5}$$

$$= \left[-\frac{1}{2}(X-\bar{\mu}_i)^T S_i^{-1}(X-\bar{\mu}_i) - \frac{N}{2}ln2\pi - \frac{1}{2}ln|S_i|\right. \tag{6}$$

2.3 Posterior Probability

Posterior probability is, when the present state of X, probability of the samples belong to different categories, this probability can be used as a basis for identifying the object ownership. Because they belong to different types of objects ,there are possibility the same values observed ,namely, the observed characteristics of a sample vector is X, while in these class have more than one class may show the value, belongs to all kinds of the probability that it be show available P(w_i/X). Can use Bayes formula to calculate the condition probability, called the posterior probability of the state:

$$P(w_i/X) = \frac{P(X/w_i)P(w_i)}{\sum_{j=1}^{n} P(X/w_j)P(w_j)} \tag{7}$$

P(w_i/X)——There is indicated the probability that the sample belong w_i class in the X conditions.

2.4 $P(w_1/X)$,$P(w_2/X)$ and $P(X/w_1)$,$P(X/w_2)$

$P(w_1/X)$ and $P(w_2/X)$ under the same condition X, comparison the probability of the appears w_1 and w_2 .If $P(w_1/X) > P(w_2/X)$,then Have the following conclusions that, the event w_1 appears the probability more than w_2 under the X condition.

$P(X/w_1)$ and $P(X/w_2)$ appear X's the probability under their conditions. No connection between the two, no meaningful comparison between the two under discussion$P(X/w_1)$ and $P(X/w_2)$ in different

3 Example of Digital Identification

3.1 Steps

(1) Prior probability calculation $P(w_i)$

It can approximation calculate through the total of sample. Here is the mumber, i=0,1,2,…,9,have,

$$P(w_i) \approx N_i/N \tag{8}$$

Where,

$P(w_i)$——is the prior probability of the digital i;

N_i——is the sample size of the digital i;

N——is total sample.

(2) Calculate $P_j(w_j)$。

$P_j(w_j)$ is express that the sample $X(x_0, x_1, x_2, \dots, x_{24})$ belong the probability estimates values of the jth is 1(x_i=1) under condition w_i.

(3) Then calculate the conditional probability $P(X/w_i)$。

The conditional probability of sample X as follows:

$$P(X/w_i) = P[X = (x_0, x_1, x_2, \dots, x_{24})/X \in w_i] \tag{9}$$

(4) Calculated posterior probability using Bayes formula.

$$P(w_i/X) = \frac{P(X/w_i)P(w_i)}{\sum_{j=9}^{n} P(X/w_j)P(w_j)} \tag{10}$$

(5) The maximum posterior probability category (0 to 9) is a handwritten digit category.

3.2 Result

Fig.2 is a sample of handwritten numbers, Fig.3 is Schematic diagram of extraction of the template, and sample library sample template example, the sample of the more recent sample, the more like or the most similar.

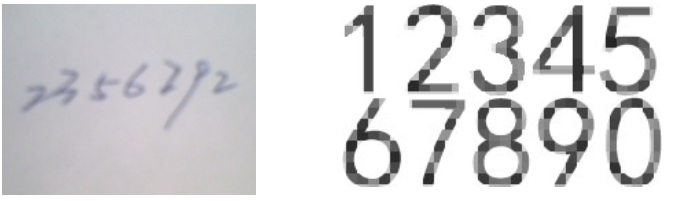


Fig. 2. Handwritten Text & Templates diagram

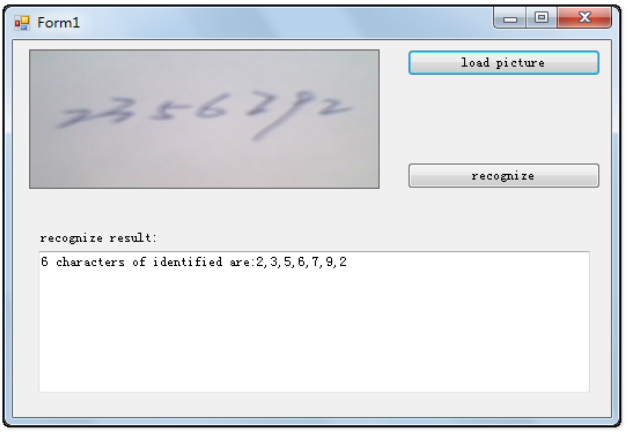


Fig. 3. Bayes classifier recognition results

Treatment using the C# programming, and standard PC. Example shows that the recognition rate of 83% or more.

4 Conclusions and Discussion

The methods used in this article only probabilistic method one of them. Making use of Bayes classification method, which the determine function and the basis for decision making are ultimately the index of similarity rate. This method has a space of application that is Vehicle License Plate recognition, Vehicle Style recognition and so on in transportation.

Acknowledgements. The Project was Supported by the Special Fund for Basic Scientific Research of Central Colleges(CHD2010JC035), Chang'an University, by the Special Fund for Basic Research Program of Chang'an University and by the open Fund for Shaanxi Road Traffic Detection and Equipment Engineering Research Center.

References

1. Jiang, Q., Li, Y., Li, W.-H., Wang, D.: Fast Face Capturing and Tracking Method in Motion Image. Journal of Jilin University (Science Edition) 47(6), 1246–1250 (2009)
2. Likforman-Sulem, L., Sigelle, M.: Recognition of degraded characters using dynamic Bayesian networks. Pattern Recognition 4, 3092–3103 (2008)
3. Toshniwal, D., Roy, R.S.: Clustering Unstructured Text Documents Using Naive Bayesian Concept and Shape Pattern Matching. IJACT 1(1), 52–63 (2009)

Imbalanced Data Problem of Relevance Vector Machine Customer Identification

Gang Li, Li Zhang, and Gui-long Wang

School of Economics &Management, Xi' an Technological University
Xi' an 710032, China
Lg21c@qq.com

Abstract. Imbalanced data problem has a significant impact on the performance of RVM pattern recognition. Customer identification is an important application domain of pattern recognition which is mapping from samples to different categories by machine learning. In order to solve the problem, the paper proposes a method named up-sampling which overcomes the phenomenon that the machine is more partial to the majority classes while ignoring the sparse and decreases the false judgment about the sparse ones.

Keywords: imbalanced data, RVM, customer identification, pattern recognition, up-sampling.

1 Introduction

Customer identification is an important application domain of pattern recognition which is mapping from samples to different categories by machine learning. Data is the ensemble of samples and provides the raw material or environment for machine learning. The performance of the machine is based on the 'understanding' and 'fitness' to the data environment. However, due to the diversity of the learning mechanism, the machines have different requirements to the quality of data, manifested in the endurance to data's noise, incompleteness, inconsistence and imbalance, etc. This research focuses on the imbalanced data problem of the Relevance Vector Machine (RVM) [1] customer identification and puts forward solutions according to the characteristics of RVM.

Imbalanced data set is that the number of negative class far outnumbers the positive, which typically occurs in practical applications. The proportion of the minority class to the majority can be drastic as 1:100[2], even approaching 1:100,000[3] in medical diagnosis [4] or direct marketing [5]. Machine learning algorithm always makes the "accuracy"[1] as evaluation criteria. Owing to the small samples occupying less part of the ensemble domain, the more of small samples even all of them are wrongly judged, which don't obviously affect accuracy. Then the classifier ignores the sparse. However, the sparse class is very important and highly concerned. Purpose of this paper is to discuss how RVM pattern recognition is sensitive to the imbalanced data and how to deal with the problem.

[1] It is the number of the samples accurately judged to the number of the ensemble domain.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 448–454, 2011.

2 RVM Inference

Given a data set $\{\mathbf{x}_n, t_n\}_{n=1}^{N}$, the target vector is $\mathbf{t} = (t_1, \cdots, t_N)^{\mathrm{T}}$, $t_n \in \{0,1\}$, which can be expressed as
the sum of an approximation vector and an 'error' vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_N)^{\mathrm{T}}$:

$$\mathbf{t} = \sum_{m=1}^{M} \omega_m \boldsymbol{\varphi}_m(\mathbf{x}) + \boldsymbol{\varepsilon} \tag{1}$$

where $\mathbf{w}$ is the parameter vector which follows a zero-mean Gaussian prior distribution:

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{m=0}^{M} N\left(w_m | 0, \alpha_m^{-1}\right) \tag{2}$$

$\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \cdots, \boldsymbol{\varphi}_M]$ is the N×M 'design' matrix whose columns comprise the complete set of M 'basis vectors'.

Unlike the regression method, the classification solution applies the logistic sigmoid link function $\sigma(\mathbf{y}) = 1/(1+e^{-\mathbf{y}})$ to $\mathbf{y}(\mathbf{x})$ and assumes that the target t_n follows the Bernoulli distribution:

$$p(\mathbf{t}|\mathbf{w}) = \prod_{n=1}^{N} \sigma\{y(\mathbf{x}_n; \mathbf{w})\}^{t_n} \left[1 - \sigma\{y(\mathbf{x}_n; \mathbf{w})\}\right]^{1-t_n} \tag{3}$$

Because the weights cannot be integrated out analytically, Laplace approximation procedure was chosen to obtain the prediction of the targets:

$$\hat{\mathbf{t}} = \boldsymbol{\Phi}\boldsymbol{\mu}_{MP} + \mathbf{B}^{-1}(1 - \mathbf{y}) \tag{4}$$

with $\boldsymbol{\mu}_{Mp} = \boldsymbol{\Sigma}\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{B}\hat{\mathbf{t}}$, $\boldsymbol{\Sigma} = \left(\boldsymbol{\Phi}^{\mathrm{T}}\mathbf{B}\boldsymbol{\Phi} + \mathbf{A}\right)^{-1}$, where $\mathbf{B}$ is a diagonal matrix $\mathbf{B} = diag(\beta_1, \beta_2, \cdots, \beta_N)$, with the element $\beta_n = \sigma\{y(x_n)\}[1 - \sigma\{y(x_n)\}]$.

3 Experiments

The imbalanced data problem of RVM is that the sparse of the positive class in the training set resulting to be ignored and wrongly judged. Take the Ripley's data set [6] for example. Ripley's data set was artificially synthetic based on the normal distribution, which is a balanced dataset, containing 625 positive samples and 625 negative ones. A Gaussian basis function with 'width' equal to 0.5 was utilized. The classification border of Ripley's data set is shown in Fig. 1.

From Fig. 1, it shows that the rate of the false judgment about the positive class is 8.00%, while the negative one is 9.60%, which is relatively balanced. The detail is in the Table 1.

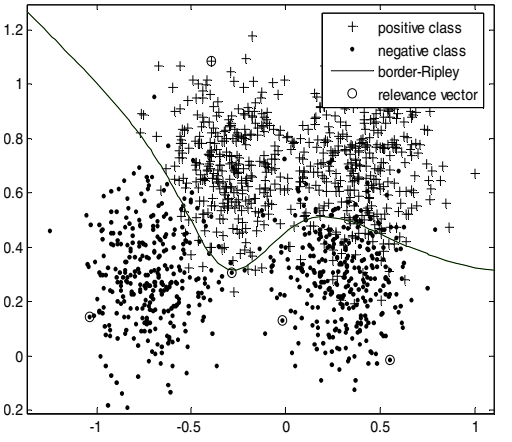


Fig. 1. RVM classification border of Ripley's data set

Table 1. RVM classification function of Ripley's data set

	Positive prediction (rate)	Negative prediction (rate)
Positive class	575 (92.00%)	50 (8.00%)
Negative class	60 (9.60%)	565 (90.40%)
Accuracy	91.20%	

Decreasing the positive class of Ripley's data set to 4%[2] of the original, which will be Ripley's positive 4% data set later. Then change the initial Ripley's data set into the imbalanced one. Use the same primary function and parameter training Ripley's data sets referred to before and get a new classification border (Fig. 2).

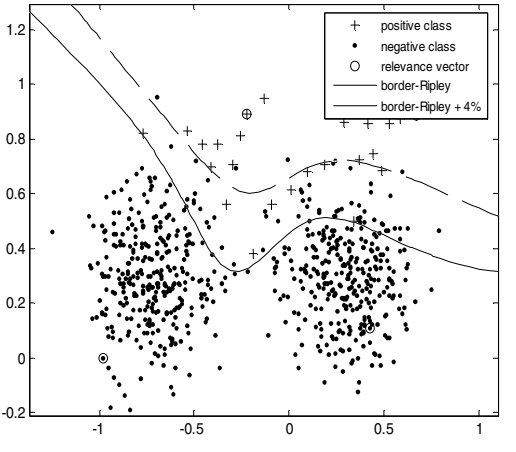


Fig. 2. RVM classification border of Ripley's positive 4% data set

[2] 4% is just to show the sparse of the positive class, with nothing else purpose.

Obviously, there's some change between the Fig.1 and Fig.2: the border of the imbalanced data moves up, then more positive classes are judged as negative ones. Furthest, the accuracy goes up from 91.20% to 98.00%,while the error rate of the positive class changes from 8.00% to 36.00%(Table 2).

Table 2. RVM classification function of Ripley positive 4% data set

	Positive prediction(rate)	Negative prediction(rate)
Positive class	16 (64.00%)	9 (36.00%)
Negative class	4 (0.64%)	621 (99.36%)
Accuracy	98.00%	

In the experiment, RVM seeks the highest accuracy of the ensemble classification, which favors the large classes while ignores the sparse ones. Namely, the sparse classes are judged as the large, then the misjudged rate of the sparse goes up.

4 Solutions of the Problem

The imbalanced data problem is mainly solved through the following ways [7-8]:

Firstly, sampling approach[9-10] is unchanging the basic classification models and using appropriate method to reconstruct the training set, which could directly pre-processes the data, then training the classifier with the trained data. There are up-sampling method and down-sampling method. The former one reduces the false judgment of the positive class through increasing the positive samples. It generally uses the method of directly copying the positive classes. The latter one reduces the false judgment of the positive class through reducing the negative samples. Compared with the latter one, the former one could not lose valuable data information. However, it would increase ensemble to enlarge the use of the classifier in space and time [11].

Secondly, algorithm method [12] is directly improving the basic classification models, without changing the samples of the training set, which adjusts the weight of the different classes to estimate the threshold and improve the models.

The first one has good independence, without the process to reform and analyze the classifier, which make it the prior method to solve the imbalanced problem. Due to the features of simple principle and easy operation about the up-sampling, we will try to use it and solve RVM imbalanced problem. Notwithstanding using Ripley's data set to illustrate the solution:

Make more copies of the positive classes and keep the total of the positive classes and negative ones essentially balance. In other words, if initial proportion of the positive to the negative is 1: n, add the positive n-1 times into the training set. As to the Ripley's positive 4% data set, add 24 positive copies into the samples and turn to Ripley's positive 4%-US data set. Use the new data set to train RVM with the same primary function and parameter ahead and get the classification border (Fig. 3).

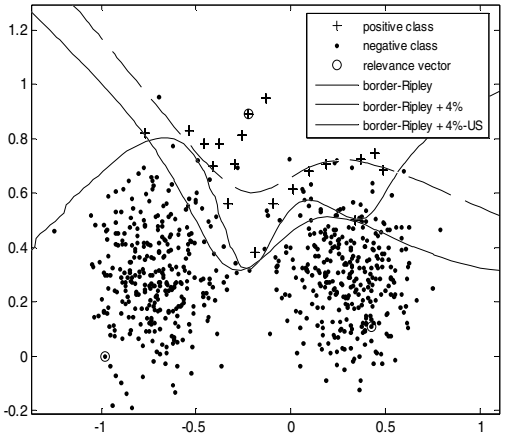


Fig. 3. RVM classification border of Ripley's positive 4%-US data set

Compared with classification border of the imbalanced data, the border of the up-sampling data obviously moves down. The range of the positive classes conspicuously expands. The attention is strengthened to the positive classes. While compared with the balance data, the up-sampling one still over the former, as the result of the information couldn't reach the number of Ripley. From the matrix of Ripley's positive 4% -US data sets and the accuracy in the Table 3, we could reach the conclusion that almost all the positive classes are exactly judged and the rate of the negative ones reaches 95.68%, while the accuracy reaches 95.84%. So the up-sampling method is suitable for the problem of RVM imbalanced class data sets.

Table 3. RVM classification function of Ripley's positive 4%-US data set

	Positive prediction (rate)	Negative prediction (rate)
Positive class	575 (92.00%)	50 (8.00%)
Negative class	60 (9.60%)	565 (90.40%)
Accuracy	91.20%	

5 Theoretical Explanation of Imbalanced Data Problem

As stated in section 2, the pattern recognition problem is equal to finding the maximum of the regularization Logistic problem:

$$\mathbf{w}_{MP} = \arg\max_{\mathbf{w}} \sum_{i=1}^{N} \left[t_n \log y_n + (1-t_n) \log (1-y_n) \right] - \frac{1}{2} \mathbf{w}^{\mathrm{T}} \mathbf{A} \mathbf{w} \tag{5}$$

with $\mathbf{A} = diag(\alpha_0, \cdots, \alpha_M)$, where $y_n = \sigma\left(y(\mathbf{x}_n; \mathbf{w})\right)$ is the confidence probability.

$\mathbf{w}_{MP}$ is the most probable vector of the weights that maximize $p(\mathbf{w}|\mathbf{t}, \boldsymbol{\alpha})$, which can be obtained when the following regularization conditions are satisfied: the samples are

properly recognized; the confidence probability y_n are closed to 1(positive class) or 0(negative class) at the most great extent; the minimum of the penalty term $\mathbf{w}^{\mathrm{T}}\mathbf{A}\mathbf{w}$ is acquired.

Because $\mathbf{w}_{MP}$ is obtained from Maximum Likelihood Estimation (MLE), the terms on the right side of (5) will increase in the iteration process. The penalty term $\mathbf{w}^{\mathrm{T}}\mathbf{A}\mathbf{w}/2$ will also increase, which can offset the increment of the Maximum Likelihood Value at some extent.

Based on the theory, the imbalanced data problem could be explained as follows:

RVM could ignore the recognition to the sparse of the positive samples. The wrong recognition of the sparse class results to the loss of the likelihood, which could be made up by the confidence probability of the negative samples. This is the direct cause to the classification border which is partial to the positive samples.

The method of up-sampling data restructure is copying the samples of the sparse class, which enlarges the proportion of the likelihood about the sparse. Then the sparse class is reached close attention, so the up-sampling method is effective on solving the imbalanced data problem.

6 Summary

In order to pursue higher accuracy of the ensemble classification, imbalanced problem brings about being partial to the majority classes while ignoring the sparse, which leads the rate of the false judgments about the sparse ones to go up. Then put forward the up-sampling method, which is suitable in dealing with imbalanced data problem in RVM learning.

Acknowledgments. The authors would like to thank the reviewers for helpful comments and suggestion. This work was supported by The National Natural Science Foundation of China (70872087) and (71002102), besides it was supported by the Program for Innovative Science and Research Team of Xi'an Technological University.

References

1. Tipping, M.E.: The Relevance Vector Machine. In: Advances in Neural Information Processing Systems, vol. 12, pp. 652–658 (2000)
2. Ling, C.X., Li, C.: Data mining for direct marketing: Problems and solutions, Citeseer (1998)
3. Pearson, R., Goney, G., Shwaber, J.: Imbalanced clustering for microarray time-series. In: Proceedings of the ICML 2003 Workshop on Learning from Imbalanced Data Sets (2003)
4. Wilk, S., Sowiski, R., Michaowski, W., et al.: Supporting triage of children with abdominal pain in the emergency room. European Journal of Operational Research 160(3), 696–709 (2005)
5. Ha, K., Cho, S., Maclachlan, D.: Response models based on bagging neural networks. Journal of Interactive Marketing 19(1), 17–30 (2005)

6. Ripley, B.D.: Pattern recognition and neural networks. Cambridge Univ. Pr., New York (1996)
7. Chawla, N.V., Japkowicz, N., Kotcz, A.: Editorial: special issue on learning from imbalanced data sets. ACM SIGKDD Explorations Newsletter 6(1), 1–6 (2004)
8. Chawla, N.V.: Data mining for imbalanced datasets: An overview. Data Mining and Knowledge Discovery Handbook, 875–886 (2010)
9. Maloof, M.: Learning when data sets are imbalanced and when costs are unequal and unknown, Citeseer (2003)
10. Kubat, M., Matwin, S.: Addressing the curse of imbalanced training sets: one-sided selection, Citeseer (1997)
11. Batista, G.E., Prati, R.C., Monard, M.C.: A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD Explorations Newsletter 6(1), 20–29 (2004)
12. Wu, G., Chang, E.Y.: KBA: kernel boundary alignment considering imbalanced data distribution. IEEE Transactions on Knowledge and Data Engineering, 786–795 (2005)

Evaluation of Sitting Comfort Based on Pressure Distribution*

Juanfang Xu

Department of Art and Design, College of Humanities,
Shanghai Sanda University, Shanghai, China
angiexjf@126.com

Abstract. In modern life, more and more work needs to be done in sitting position. So sitting comfort is of vital importance. However, comfort was always different by personal feelings. This paper studied the evaluation of sitting comfort in the way of comparing subjective comfort scales and objective computer data, which was carried out by the equipment of X-sensor pressure distribution system. Through two experiments and with formulas analysis, materials of chairs were evaluated of their comfort characteristics.

Keywords: body pressure distribution, sitting comfort, X-sensor system.

1 Introduction

Nowadays, more and more works should be carried out in sitting postures. People get tired sitting for a long time. chronic diseases such as cervical spondylosis and lumbar became so common due to long-term sitting. Sitting comfort is worth a major concern and subject of study. "Comfort level" is usually defined as the personal satisfaction in a particular state of the environment. Currently, the research of comfort measurement is generally in two ways, that is, from objective and subjective aspects. Subjective evaluation using the semantic differential method divided into ten scales. Objective measuring of comfort mainly refers to the corresponding data obtained through a variety of instruments to measure various indicators of the human body to infer human comfort. In this paper, body pressure distribution data of sitting was collected by experiment with the equipment of X-sensor pressure distribution system. And the relationship between subjective comfort feeling and the data obtained was studied.

2 Comfort Levels

Pressure distributions under different sitting postures were not the same. According to the adaptation of the human body, "comfort zone" can be divided into six parts:

Comfort zone (C): various indicators of normal, people in this position under a long time to accept and not feel excitement and fatigue;

* Sponsored by Scientific Training Selected Outstanding Young Teachers Special Fund in Shanghai Universities (ssy08035).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 455–459, 2011.

General comfort zone (U): the individual condition indicators deviated from the normal comfortable, but people can accept this position under, and "subjects" no "not comfortable" expression;

No comfort zone (A): an indicator of several conditions, the normal deviation from a comfortable, long time in this environment, "subjects" to be "not comfortable" expression;

Fatigue zone (F): in such an environment people will feel irritation and fatigue, but not pain;

Can not stand area (P): also known as "pain zone", a complete departure from the comfort of several indicators of normal conditions, people will feel the pain, unbearable.

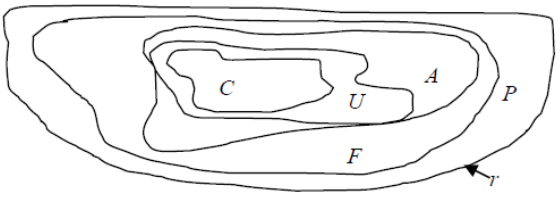


Fig. 1. Comfort Zone

3 Experiment

The first step of sitting pressure distribution comfort experiment was to measure the pressure distribution data by one subject sitting on different material chairs. In the meanwhile, asked the volunteer about the feeling of sitting comfort, and then compared the objective data and subjective comfort numerical. Through the combination study of objective and subjective experience approach, the comfortable seating surface pressure distribution could be described. The second step of the experiment was to make different subjects in the same seat, test the pressure distribution, and detect the reliability of the data.

3.1 Equipment

The experiment used the body pressure test equipment: XSENSOR pressure distribution tester. The instrument through which pressure sensors on the human body and the contact surface pressure distribution generated by the test, which includes three flexible pressure sensor test pad, respectively, for the back, buttocks and to test the human body lying down. Three sensory test pad through the small square in the grid cell for stress measurement test, the output of three sensor pads were 40 × 40,40 × 64 and 64 × 160. Pressure sensor installed in each small box inside the body using capacitive techniques to measure the pressure points, sampling port into the computer by using software to record supporting XSENSOR. This pressure distribution tester is widely used now, you can contact surface for a variety of stress tests.

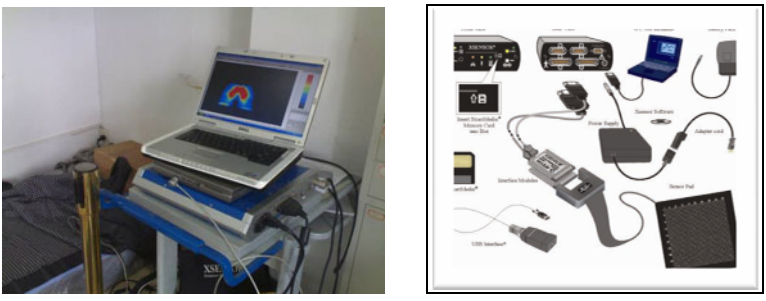


Fig. 2. XSENSOR Pressure Distribution Tester

3.2 Experimental Content

Experiment name: Comparing of different material seats in subjective sitting comfort and pressure distribution tests

Experimental equipment: pressure distribution pad, wooden chairs, plastic chairs, leather seats, memory foam seat, no back

Subject volunteers: 30 subjects (of which 20 were 20-40 years old, 10 were 40-60 years old, divided equally between men and women)

Experiment time: 2 hours each person

Experimental steps:

0. Introducing the experiment contents and steps;
1. Select seat A, the subject filled out the current scale of subjective comfort
2. Measure the pressure distribution record under stable condition of sitting
3. Asked the subject remain seated for some time (1 minute, 5 minutes) and again recorded the scale of subjective comfort
4. Standing up and rested for 1 or 2 minutes
5. Replacement of Seat B (C / D /...) and repeat 1 to 4
6. To pay fees

3.3 Experimental Results

Figure 3 for the 30 seats in various subjects through one by one experiment, analysis XSENSOR software data, and each person averaged pressure distribution data, obtained in the XSENSOR software pressure distribution comparison. From left to right in accordance with, in order from top to bottom order: wooden chairs, plastic chairs, leather chairs, foam-padded seats.

The higher the pressure parts of the figure, the more color tended to red. In the opposite, the lower the pressure parts of the figure, the less color trended to red and more color trended to blue. From the pressure distribution, we could see on the chart that wooden chair up red area, and the transition boundary was obviously. And the pressure distribution covered the maximum span. In contrast, leather seats and foam padded seat almost had no red zone. The pressure maximum was relatively small, and

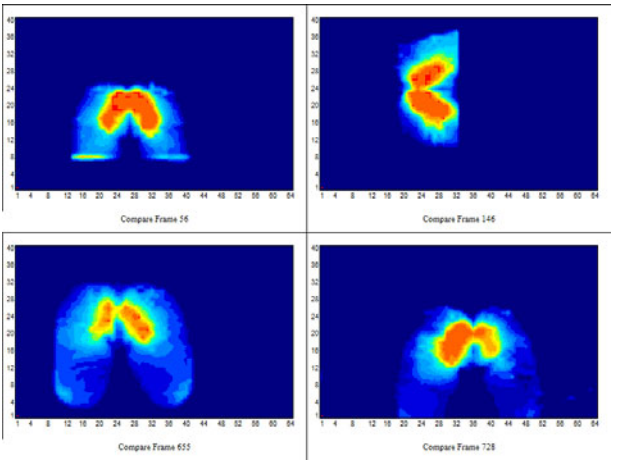


Fig. 3. Pressure Distribution of Different Materials Seats

the transition boundary was ease, indicating a more uniform pressure distribution. And the local pressure was relatively balanced. The same conclusion can be based on the following formulas.

$$p_m = \max(p_1, p_2, \cdots, p_n) \quad \textbf{(1)}\ \text{Max}$$

$$p_v = \frac{1}{N_p}\sum_{i=1}^{N_p} p_i \quad \textbf{(2)}\ \text{Average}$$

$$G_m = \max(\text{grad}\,G_1, \text{grad}\,G_2, \cdots, \text{grad}\,G_n) \quad \textbf{(3)}\ \text{Maximum Gradient}$$

$$G_v = \frac{1}{N_p}\sum_{i=0}^{N_p}(\text{grad}\,G_i) \quad \textbf{(4)}\ \text{Average Gradient}$$

In addition, comfort is a subjective test scale divided into 10 grades. The most comfortable was 10, followed by decreasing according to comfort, the least comfortable was 1. According to the results of subjective comfort, the average value calculated by the average of four comfortable seats were: 5.6, 6.3, 7.5, and 7.4.

Compare the data based on subjective comfort and the pressure distribution, two results basically the same conclusion. It's not hard to expose the fact that foam-padded seat and leather seat were more comfortable than wood and plastic seats. And of the two the plastic seat was more comfortable than the wood one.

The second step was to test for the same seat, with the pressure distribution of different subjects. Which were compared to determine the different people in the same seat was able to show the same characteristics.

From Fig.4, Wooden seat pressure distribution was more concentrated, the transition boundary was obvious, while the foam-padded chair was moderate. So the pressure distribution was relatively uniform.

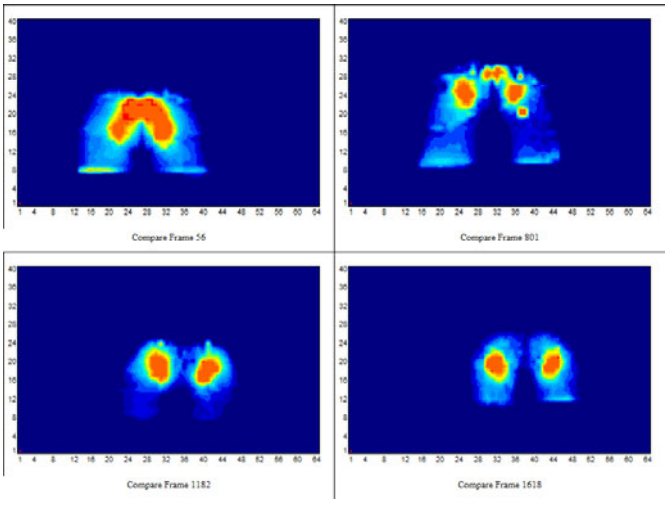


Fig. 4. Pressure Distribution of Different Subjects in the Same Seat(wood)

4 Conclusion

According to the above experiment, leather and foam padded seats pressure reflected in the body pressure distribution were closer to the optimum. So the two materials were more comfortable. From the experiment results, we can infer that, of the four kinds of materials, soft ones were more comfortable than the hard seat surface. The soft material can be more appropriate than the hard surface to fit the hips. The pressure distribution was more balanced, that muscle fatigue caused by long-sitting position could be carried down. Thus by comparing pressure distribution and subject comfort feelings, evaluation of sitting comfort can be done.

References

1. Xiao, Y.: Seat Comfort and Ergonomics. Railway Vehicle 35(5), 32–34 (1997)
2. Chen, R.: Based on the work of maintaining a healthy sitting position of chair design. Master Thesis, Zhejiang University (2007)
3. Gruenberg, M.F., Rechtine, G.R., Chrin, A.M., Sola, C.A., Ortolan, E.G.: Over distraction of cervical spine injuries with the use of skull traction: a report of two cases. J. Trauma 42, 1152–1156 (1997)
4. Liu, X., Dong, Z.: Pressure distribution based on experimental data, the product form of human contact with surface design. Computer Integrated Manufacturing System 14(4), 644–646
5. Joines, S.M., Sommerich, C.M., Mirka, G.A., Wilson, J.R., Moon, S.D.: Low-level exertions of the neck musculature: a study of research methods. Journal of Electromyography and Kinesiology 16, 485–497 (2006)

The Research on Human-Computer Interaction of Remote Disaster-Recovery System

Wang Chao, Li Yuanbin, and Duan Xuliang

School of Information and Engineering Technology, Sichuan Agricultural University, Ya'an Sichuan 625014, China

Abstract. The human-computer interaction(HCI) technology was researched and applied, which makes the interaction system has full functions, immediate response, simple and smooth operation, and improves user experience. And as a result of getting rid of redundant data in interaction process to a great extent, this solution greatly brings down the traffic flows between the client-side and the server-side, and relieves the burden and bandwidth stress on the server-side.

Keywords: Remote Disaster-Recovery System; Human-computer Interaction; Interaction process.

1 Introduction

Along with the rapid development of information technology, computer-information system has become an important property of companies. As the system is playing an important role in enterprise's operation, the security is becoming more and more indispensable [1].

Remote Disaster-Recovery System is generally considered as an effective solution for improving system available and resisting large-scale disasters. Remote Disaster-Recovery System concerns the local-remote data changes and networks, as it is complex in operation and difficult in management, the developing of a fully functional, quality and stability, user-friendly Disaster-Recovery System is very necessary. However, in traditional Disaster-Recovery mode, weak client function and huge redundancy communication data limits the system application and promotion.

This paper is to discuss the above topics with human-computer interaction technology. Based on this technology, the Remote Disaster-Recovery System maximize eliminated the interaction decrease data, reduced the interaction communication flow and greatly alleviate the server strain and bandwidth pressure.

2 Topological of Remote Disaster-Recovery System

The Remote Disaster-Recovery System introduced is an Internet-based system which consists of two parts: local-datacenter and remote recovery datacenter, as shown on Figure 1.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 460–466, 2011.

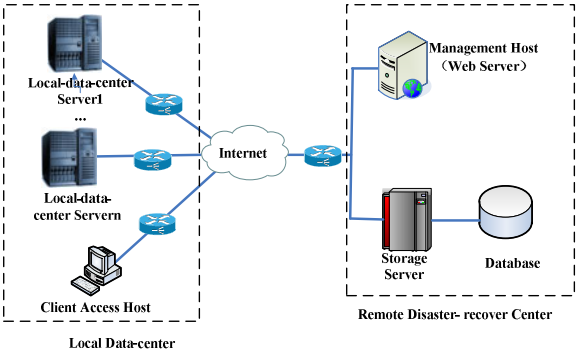


Fig. 1. Topological of Remote Disaster-Recovery System

Local datacenter is composed of several local-servers which are applied to backup and extra clients for the purpose of logging on and operation.

The manager-hosting and storage-server located in the remote disaster-recovery center of the same LAN and large communication bandwidth supporting for fast access to system databases. The host is using for system configuration-management while the storage server receive and process in-time backup-data.

3 HCI Technology

Human-Computer Interaction (HCI) is a technology of human, computers and their mutual influence, which is considered as the communication medium and interface. The core functional requirements must be considered in the designing of the HCI system such as user positioning and the choice of interaction patterns. In short, there are several factors that would influence human-computer interaction system design as following:

Generally, user &flow compatibility, consistency, flexibility, error handling ability are the basic principle of system designing. [2]

User Interface Management System (UIMS) is the program framework of client software interface, which theoretically describes the working principle of human-computer interaction systems, structure and human-computer interaction. Widely used in UIMS is the SEEHEIM model [3], that the basic idea is to separate the user interface and applications, dialogue and application linked by a control unit. In SEEHEIM model, the three parts of user interface is shown in Figure 2.

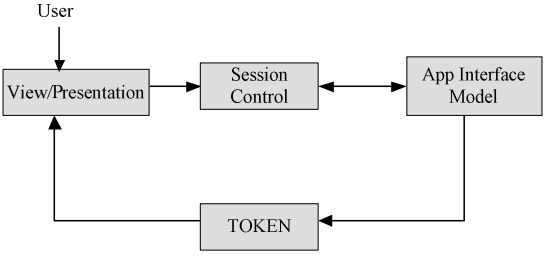


Fig. 2. User-Interface of SEEHEIM model

4 HCI Design

The human computer interaction designing is rooted in the functions of remote disaster recovery. Core function of the system is to backup client-data to the remote recover center, and to restore client-data while destroyed. All the operations associated with the client user in the system are the contents of human computer interaction design.

The topology of system (Figure 1) shows that the host and client-server are in the same subnet of a local data center, which two connected by Internet. The interaction design of the system shows in Figure 3:

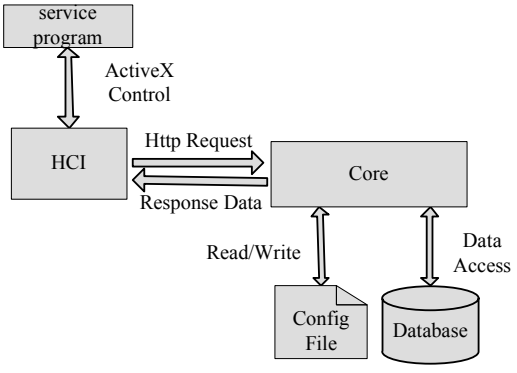


Fig. 3. Schematic diagram of HCI

The system is designed as needed for operations. Corresponding operation commands such as start/stop/initialization etc. is sending to client-server while the service programs running on the client-server receiving and handling these commands. Since the local data center may be in a private subnet, commands only can only be sending by the access-host browser instead of the management-server in remote data center.

As in Figure 3-1, browser communicates with client-server service program through ActiveX controls. The ActiveX control defines certain functions methods to communicate with the service program. Embedded in human-computer interaction page, ActiveX control is called by browser program, and establishes a TCP connection with client-server services process, run custom control programs, such as send commands to the client-server service program and returns operation results.

The server response information is divided into two forms: plain text messages and XML documents. Text messages are easier to rapid resolving for human-computer interaction engine. For the simpler applications that require rapid response and with small communication, plain text could reduce system processing complexity, speed up the response and processing. The superior hierarchical structure of XML document is better use for a large-scale data, or with complex structures. The XML can also be used as common communication format between heterogeneous applications. Owing to traversing all XML nodes, parsing XML taxing the system and the processing is much slower than that of plain text. However, the interaction engine communication data

between clients' browser and server is small and simple-structured. As it is the requirement of rapid response plain text is rather suitable than XML.

Messaging protocol for server response to the HCI engine compose of message type, separator, task data and terminator.

5 Implementation of HCI

A . Main functions

The main function involved in HCI of remote disaster recovery system is described as follows:

TaskRequest ();

Function: Create asynchronous requests of obtaining task run info on client, and send to server;

Implementation: The function is triggered by Window.onload Event of the page, and then calls HCI engine to create an asynchronous request. XMLHttpReques() is called to create XMLHttpRequest object as following codes:

xmlhttp.open ("GET", "task.jsp", true);

xmlhttp is the variable of the object values XMLHttpRequest, taskinfo.jsp is a server-side file of handling disaster task.

Note that, when access task.jsp the cache-control mode must be set, since this function is called by other functions to send asynchronous requests regularly, task.jsp is accessed many times. Browser may take optimization when access the same server file: Cached the file on client as first visit, and later, directly access the file from the cache instead of retrieving from server. This optimization would speed up access rate and reduce the downloaded data and especially effective for the page that not always changes. However, it would cause trouble for the file which feedback the latest running information of disaster recovery tasks after each access; the browser would always obtain the response of first access the file. So the file must be set to no-cache, browser every access of the file will be from the server instead of using client cache. Settings of task.jsp cache control mode of are as follow:

header("Cache-control: No-Cache");

taskRespone();

Function: Clients function of processing server response.

Implementation: the function receives server response information, obtains and updates the page fields' value. After processing the response from the server, 3-second intervals then the function call TaskRequest() function again and send new asynchronous requests to obtain the latest operation information.

B. Message Protocol

Server gets asynchronous request that the latest run information of task, response message format is used: " human-computer interaction _MONITOR | task | bdstatus | netstatus | speed | waitdata | percent | runstastus | datastatus \ n ", where "human-computer interaction _ MONITOR" is the message type, "task" is the task, "bdstatus, netstatus, speed, waitdata, percent, runstastus and datastatus" used to update

task information of the page which are client-server status, network connectivity status, data backup flow, the amount of data is not mirrored, the data mirroring percentage, task status and data status.

6 System Performance Analysis and Testing

Based on HCI technology, the remote disaster recovery system is easy to use as a desktop application. More important, interactive technology greatly reduced the network communications between browser and server, effectively eased the server bandwidth pressure and response stress.

Testing of the remote disaster recovery system is in a 100Mbps network. Test environment shown in Table 1.

Table 1. Testing environment

	CPU	*memory*	*harddisk*	*OS*
client	Intel® P41.6GHz	256M	80G	Win XP
server	Intel® Celeron(R) .40GHz	512M	160G	Linux2.4.31

In order to prove that the HCI improve system performance, in the system network flow testing we will use peer-to-peer system to compare interactive mode with traditional, and analysis system performance according to the test data.

Continuous monitoring of browser on the two comparison systems lasts 6 minutes, the sending and receiving data are recorded once a minute. Statistics of network traffic data are shown in Table 2 (units: M):

Table 2. The network traffic monitoring data of traditional and human-computer interaction model in the remote disaster recovery system

time	traditional mode			Interaction mode		
	send	receive	total	send	receive	total
1	104.6	134.3	238.9	9.6	39.2	48.8
2	198.3	254.4	452.7	17.8	52.6	70.4
3	297.4	381.6	679.0	26.2	66.3	92.5
4	394.0	507.8	901.9	34.6	80.2	114.8
5	490.2	628.9	1119.1	42.8	93.6	136.3
6	589.3	756.1	1345.4	50.5	106.3	156.8

Here, send and receive columns represent the sending and receive traffic, total is the total network traffic of sending and receiving. Each row of data line shows the total

network flow within n minute of the two contrasting systems. Such as the first line means the total network flow within one minute, second line means the total network flow within two minute, and so on.

Collation and statistical data in the table, in total line use each row subtract the row before, obtained six groups network flow within one minute, as shown in Table 3 (units: M):

Table 3. Five group network traffic statistics and comparison within one minute of two modes

Time	traditional	interaction	multiple
1	213.8	21.6	9.9
2	226.3	22.1	10.2
3	222.8	22.2	10.0
4	217.2	21.6	10.1
5	226.3	20.5	11.1

As it is shown in table, column "multiple" is the traditional model multiple of interactive mode; it's evident that former value is about 10 times than the latter. In other words, use HCI model can reduce the 90% of network traffic that in traditional model. The improvement of system performance is very obvious.

The test data show that compared with traditional model of remote disaster recovery system, based on HCI model could reduce redundant requests, ease server stress, and diminish the download data. As decreased in communication information of client and server, the server bandwidth stress is significantly improved.

7 Conclusions

The remote disaster recovery system with human-computer interaction technology has a series characteristic such as a friendly interface, easy operation, complete in functions and superior performance. It as an important component of the remote disaster recovery system and would help to promote the use of remote disaster recovery system.

References

1. Li, T.: Introduction to Network security, 1st edn., pp. 474–490. Electronic Industry Press, Beijing (2004)
2. Mayhew, D.J.: Principles and guidelines in software user interface design. Prentice-Hall, Englewood Cliffs (1992)
3. Huang, B., Zhao, Z.: User Interface Management System Design and Construction. Computer Engineering and Design 5 (1993)
4. Sun, J., Liu, X., Li, T., et al.: The design and implementation of Internet-based remote disaster-recovery system. Computer Applications 24(7), 171–172 (2007)
5. Sha, J., Liu, X., Li, T., et al.: A remote control based on security system design and implementation of disaster recovery. Computer Applications 24(8), 152–154 (2007)

6. Wang, K., Zhou, L., Cai, Z., et al.: A disaster recovery system model in an e-government system. In: Proc. of the 6th International Conference Parallel and Distributed Computing. Applications and Technologies, pp. 247–250 (2005)
7. Source, A.: SANs sharing storage to infinity and beyond. Electronic Design 52(9), 50–59 (2004)
8. Liu, L.: SCM security solution based on SSL protocol. In: IEEE International Conference on Service Operations and Logistics, and Informatics, SOLI 2006, pp. 814–817 (2006)
9. Michener, J.R., Acar, T.: Managing system and active-content integrity. Computer 33(7), 108–110 (2000)

An Experimental Design for Interrupt Controller of Embedded System

Zhang Donghui

Department of Computer Engineering, Zhongshan Polytechnic, Zhongshan 528404, China
zhangdhok@126.com

Abstract. In most case, microprocessors might spend much time on testing the I/O port or component state, while adopting interrupt operation can overcome the shortcoming of CPU inquiring and improve the waiting efficiency of CPU. Interrupt handling can be classified into two categories: hardware interrupt and software interrupt. Adopting interrupt controller in synchronization operation can realize the concurrent work of CPU and peripherals and assure the computer has real-time process and control ability. Therefore, it is necessary to arrange a proper experiment to study the behaviour of interrupt controller. The interrupt handling process of S3C2410X is described, normally, S3C2410X interrupt controller requests interrupt from kernel interrupt FIQ or IRQ of S3C2410X after the interrupt arbitration. The experiment purpose and experiment equipment is detailed listed. Finally, the author explored into the experiment principle and experiment procedure.

Keywords: Experimental design, Interrupt controller, Mode.

1 Introduction

For an embedded system, the microprocessor that controls the I/O port or data transferring modes can be classified as three kinds, which are procedures inquire mode, interrupt mode and the DMA mode, respectively. Procedures inquire is a microprocessor lookup routine that periodically inquires the program to read I/O port or content of status register and judges its state. In most case, microprocessors might spend much time on testing the I/O port or component state, and the data of port or components also cannot get real-time processing, therefore, the efficiency of program inquire mode for a microprocessor normally is very low. When a port or component required to transmit data, the data cannot transmit to target before microprocessor inquires into the program. Adoption of interrupt operation can overcome the shortcoming of CPU inquiring and improve the waiting efficiency of CPU, thus it had been widely applied in embedded system.

In the past days, many scientist and engineers have paid much attention to the fields interrupt contorller. Yves Blaquière in University of Quebec examined the procedures and used of hardware and software tools in the design and simulation of the circuit[1]. Stergios Spanos in National Technical University of Athens analyzed the routing performance improvement of an embedded communications processor by

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 467–474, 2011.

endorsing advanced interrupt handling techniques in 2008[2]. Phillip John McKerrow in University of Wollongong used of a logic-state analyser as a hardware monitoring tool, in his study, the common interrupt-handling program of the UNIX operating system, running on a Perkin-Elmer 7/32, was measured before and after major code modification[3].

Interrupt handling is a communication mode happens between external devices, programs and CPU. As soon as CPU received interrupt signal, the system will suspend the running program at the same time and turns to process the interruption task, and this interrupt handling program is called interrupt service routine. When the interrupt service routine is finished, CPU will return to original interrupted procedure. Interrupt handling can be classified into two categories: hardware interrupt and software interrupt. Hardware interrupts are produced by external hardware, which can be divided into two classes: the shielding interrupt and non-inhibit interrupt. Software interrupt has no relationship with hardware circuit, because it is produced by the setting of registers or CPU instruction[4].

The adopting of interrupt controller in synchronization operation can realize the concurrent work of CPU and peripherals, which can greatly improve the utilization rate of CPU and increase the output speed. Meanwhile, the real-time processing can timely handling various parameters and the information that randomly loading into computer, which can assure the computer has real-time process and control ability. The most important mission for interrupt handling is resolving fault. CPU can send an interrupt request to execute the corresponding fault handling procedure according to the fault source immediately, dispose of fault without stopping, by which the working reliability of the computer is improved.

2 Experiment Purpose

The experiment purpose consists of five aspects.

(1)Master the operation of interrupt control register S3C2410X by experiment

(2)Master the response process of interrupt control register S3C2410X by experiment

(3)Master the detailed process of interrupt under different triggered methods

(4)Master the interrupt modes and interrupt process of AMR processor by experiment

(5)Master the programme composition method of interrupt process by experiment

3 Experimental Equipment

The basic experiment equipment used for this study is listed below.

(1)Hardware: Embest Edukit-Ⅲ experimental platform, ULINK2 simulator, computer.

(2)Software: μVisionIDE. Windows 98/2000 / NT/XP.

4 Experiment Content

Write interrupt service routine to realize the following functions.

Conduct external interruption EINT0, EINT11 by choosing interrupt triggered modes

Press the button of SB1202 under different triggered modes to trigger EINT0 in the experiment platform of Embest Edukit-Ⅲ, at the same time, displays external interrupt signal in super terminal of main window.

Press the button of SB1203 under different triggered modes to trigger EINT11 in experiment platform of Embest Edukit-Ⅲ, at the same time, displays external interrupt signal in super terminal of main window.

5 Experiment Principle

S3C2410X interrupt controller can accept up to 56 interrupt sources of the interrupt request. The interrupt source S3C2410X could be provided by inside or outside of chips, such as DMA, UART and ⅡC etc, among which UARTn interrupter and EINTn interrupter are logical relationship of "or". They shared with same interrupt request line.

The interrupt source S3C2410Xalso can be provided by input pins of external interruption in processor, the interrupt source are shown below.

(1) INT_ADC: A/D conversion interrupt.
(2) INT_TC: touch screen interrupt.
(3) INT_ERR2: UART2 receive and dispatch error interrupt.
(4) INT_TXD2: UART2 send interrupt.
(5) INT_RXD2: UART2 receive interrupt.
(6) INT_ERR1: UART1 receive and dispatch error interrupt.
(7) INT_TXD1: UART1 send interrupts.
(8) INT_RXD1: UART1 receive interrupt.
(9) INT_ERR0: UART0 receive and dispatch error interrupt.
(10) INT_TXD0: UART0 send interrupt.
(11) INT_RXD0: UART0 receive interrupt.

Interrupt controller S3C2410X requests interrupt from kernel interrupt FIQ or IRQ of S3C2410X after the interrupt arbitration. Interrupt arbitration process relies on hardware priority logic of the processor, at the end of arbitration process, the processor would records arbitration outcome to INTPND register, the user also have been informed which interrupt source generated the interrupt. The processing process of S3C2410X interrupt controller is shown in Fig.1.

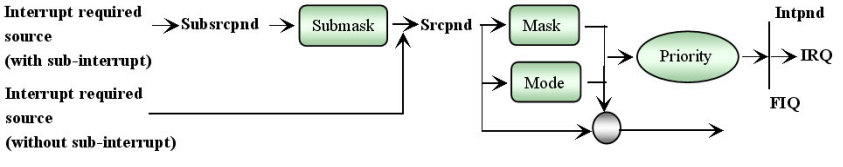


Fig. 1. Interrupt controller of S3X2410X

The task of S3C2410X interrupt controller is that, when more than one interrupts occurs, select one and send out an interrupt request from IPQ or FIQ to CPU core.

In fact, the first CPU core only have two interrupts, which are FIQ (Fast Interrupt Request) and IRQ (Interrupt Request), other interrupts are the extended definition by adding an interrupt controller when design each chip, the running process of these interruptions are instructed by interrupt priority level which are more correspond with practical applications of requring multiple interrupt sources. For example, if you define all interrupt sources as IRQ interrupt (set by the interrupt mode register), and there are 10 interrupts sending out request at the same time, then you can determine which priority interrupt should be executed by reading the interrupt priority registers.

The interrupt handling process of S3C2410X can be described as follow: When an interrupt source send out interrupt request, the interrupt controller will handle the interrupt request. According to handling results, sent IPQ request or FIQ request to the CPU core, at the same time, program counter (PC) of CPU point to the exceptional entry (OX18) of IPQ or exceptional entry (ox1c) of FIQ, and then the procedures start to perform from the exceptional entry (OX18) of IPQ or exceptional entry (ox1c) of FIQ.

External interrupt EINT4~7 and EINT8~23 are two external interrupt by the expansion of interrupt management function of S3C2410 IO register. This section will cover the usage of these registers later.

The function introduction of S3C2410X interrupt control is shown below.

(1)Position F and position I of current program status register

If position F of CPSR is set as 1, CPU will not accept FIQ from interrupt controller. If position I of CPSR is set as 1, CPU will not accept IRQ from interrupt controller. Therefore, in order to realize function of FIQ and IRQ, position F and position I of CPSR must be cleared. Moreover, corresponding position of interrupt mask register must be cleared.

(2) Interrupt mode (INTMOD)

ARM920T provides two interrupt modes, namely FIQ mode and IRQ mode. All interrupt sources need decide which interrupt mode to choose during interrupt request.

(3)INTPND

S3C2410X has two INTPND: SRCPND and INTPND, which could be used to indicate if corresponding interrupt is activated. When interrupt source makes interrupt request, the corresponding position is set as 1. Meanwhile, only one position of INTPND is automatically set as 1 after arbitration procedure. If mask bit is set as 1, corresponding SRCPND position is set as 1, but INTPND register will not change. If INTPNd is set, the corresponding interrupt service routine will be executed as long as sign I and sign F is to be reset. The interrupt service subroutine sets corresponding position of SRCPND as 1 so that source hung state will be cleared. By the same way, the hung state of INTPND corresponding position is cleared. "INTPND= INTPND" may clear lest import of incorrect data causes error

(4)INTMSK

If mask bit of INTMSK is 1, corresponding interrupt will be forbidden. If the mask bit of INTMSK is 0, corresponding interrupt is to be executed normally. If mask bit of one interrupt is 1, suspend will be set as 1 when the interrupt make request, but interrupt request will not be accepted.

Among 56 interrupt sources, there are 30 supply interrupt controller. In these interrupt sources, the external interrupt EINT4/5/6/7 supplies interrupt controller through "logic or" format. For EINT8-EINT23, it also could work in same way. Interrupt source of S3C2410X is shown in Table 1.

Table 1. Interrupt source of S3C2410X

Interrupt source	Function	Interrupt arbitration group
INT_ADC	ADC BOC and Touch interrupt	ARB5
INT_RTC	RTC alarm interrupt	ARB5
INT_SPU	SPU interrupt	ARB5
INT_UART0	UARTO interrupt	ARB5
INT_ⅡC	ⅡC interrupt	ARB4
INT_USBH	USB host interrupt	ARB4
INT_USBD	USB Device interrupt	ARB4
Reserved	Reserved	ARB4
INT_SPI0	SPI0 interrupt	ARB4
INT_SDI	SPI interrupt	ARB3
INT_DMA3	DMA channel 3 interrupt	ARB3
INT_DMA2	DMA channel 2 interrupt	ARB3
INT_DMA1	DMA channel 1 interrupt	ARB3
INT_DMA0	DMA channel 0 interrupt	ARB3
INT_LCD	LCD interrupt	ARB3
INT_UART2	UART2 interrupt	ARB2
INT_TIMER4	Timer4 interrupt	ARB2
INT_TIMER3	Timer3 interrupt	ARB2
INT_TIMER2	Timer2 interrupt	ARB2
INT_TIMER1	Timer1 interrupt	ARB2
INT_TIMER0	Timer0 interrupt	ARB2
INT_WDT	Watch-dog timer interrupt	ARB1
INT_TICK	RTC Time tick interrupt	ARB1
nBATT_FLT	Battery fault interrupt	ARB1
Reserved	Reserved	ARB1
EINT8_23	External interrupt 8~23	ARB1
EINT4_7	External interrupt 4~7	ARB1
EINT3	External interrupt 3	ARB0
EINT2	External interrupt 2	ARB0
EINT1	External interrupt 1	ARB0
EINT0	External interrupt 0	ARB0

Interrupt controller of S3C2410X have 5 control registers: SRCPND, INTMOD, INTMSK, PRIORITY, INTPND. Sent by interrupt source, interrupt request is first put into SRCPND. INTMOD sorts interrupt requests into two groups: FIQ and IRQ, PRIORITY.

6 Circuit Principle

In this experiment, the selected external interrupts are EXTINT0 and EXTIN11. Interrupts are generated by pressing button SB1202 and SB1203, when one button is pressed, EXTINT0 or EXTINT11 will come to earth connection condition and input low level, consequently system will send out interrupt request to CPU. After receiving interrupt, the CPU come into appropriate interrupt service procedure, show the current interrupt number via the main window of hyper terminal. The circuit diagram is shown in Fig.2.

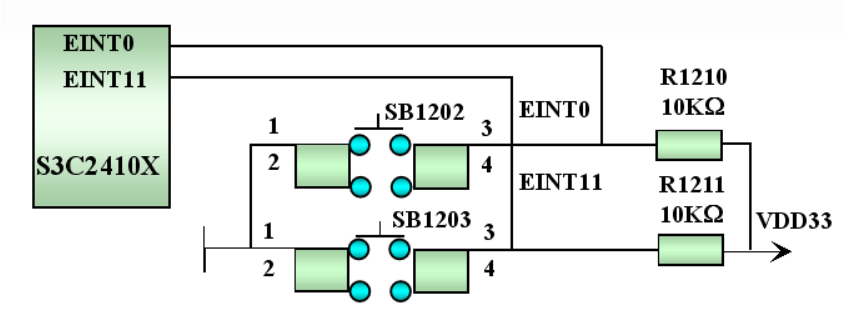


Fig. 2. Circuit diagram of interrupt experiment

7 Experimental Procedure

7.1 Prepare Experimental Environment

The experiment environment should be prepared firstly.

(1) Copy the content of folder of Code \ Chapter6 from CD-ROM to the host (if it have been copied, then skip).

(2) Use serial cable of EduKit-Ⅲ target version connect UARTO of target version and serial port COMXPC, then assembly kit of emulator ULINK2.

7.2 Setup Serial Receiver

Run hyper terminal serial communications program of windows (baud rate is 115 200Bd, the first location is stop bit, invalid check digit, no control of hardware flow), or use other serial communication program.

7.3 Operation Flowchart

The operation flowchart is listed below.

(1) Open environment of μ Vision IDE, enter into the directory of experimental routine Edukit-Ⅲ 2410\int _test, compile link works.

(2) Click the Debug menu of IDE, select the item of "Debug→Start/Stop Debug Session" or press combination key of Ctrl + F5, connect to target version and start the debugger.

(3) Click the item of interrupt controller in peripherals menu, open the window of interrupt controller, observe the value change of interrupt control register during experiments.

(4) Double-click int_tes.c in project management window to open the file, set breakpoint at the first statement of two interrupt service procedure, respectively. The breakpoint is set at the statement of function call "int_tes. ();", select the menu Debug→Run or press key F5 to run program, if program runs correctly, hyper terminal would show following information:

(5) As the program is broken down, use PC keyboard input the setting interrupt trigger mode (default 3), and then press key F10 in program interface. Observe the value of interrupt control registers, namely interrupt configuration. If choose 2, input hi pot will keep on interrupted condition.

(6) Continuously press key F10, as the cursor run to (while (g_nKeyPress & (g_nKeyPress <6))), it will keep on waiting for an interrupt from button. Press the button of SB1202 or SB1203 and press key F10 for twice, the program will break at the break entry point of interrupt service routine. Observe the value of interrupt control registers again, right-click INTERUPI, refresh register window, observe the change of each value before and after the program running (the interrupt request flag bit should be set).

(7) Remove the breakpoint, download the execution program again, press number keys to select appropriate interrupt trigger mode, press the button SB1202 or SB1203, observe that whether the output is consistent with the fact in hyper terminal main window.

(8) Select the different external interrupt trigger mode, observe the output interrupt situation of the hyper terminal when press the button SB1202 or SB1203 under different interrupt trigger.

(9) Integrated with experimental contents and experimental principle, master the interrupt operation of ARM processor, such as activating interrupt, setting trigger mode of the interrupt and identifying interrupt source, the focus area is understanding the process of terminal response and interrupt handling in ARM processor.

8 Summary

Interrupt handling is a communication mode happens between external devices, programs and CPU. The author summarized the application of interrupt handling and planned to build a proper experiment to study the behaviour of interrupt controller. The interrupt handling process of S3C2410X is described, in this study, the experiment purpose and experiment equipment is detailed listed. Finally, the author explored into the experiment principle and experiment procedure.

References

1. Blaquière, Y., Davidson, J.: VHDL design of a priority interrupt controller and synchronizer for the MC68008. Microprocessors and Microsystems 14, P474–P478 (1990)
2. Spanos, S., Meliones, A., Stassinopoulos, G.: The internals of advanced interrupt handling techniques: Performance optimization of an embedded Linux network interface. Computer Communications 31, P3460–P3468 (2008)
3. McKerrow, P.J.: Evaluation of interrupt-handling routines with a logic-state analyser. Performance Evaluation 3, 277–288 (1983)
4. Liu, H.: ARM embedded system structure and interface technology. China Posts and Telecom press (2009)

Analysis on Tort Liability of Internet Service Providers

Xiaomei Hong[1,*] and Fengyun Li[2]

[1] School of Humanities and Law, Northeastern University, Shenyang 110819, P.R. China
hongxiaomei@sina.com
[2] Computer Center, Northeastern University, Shenyang 110819, P.R. China
lifengyun@cc.neu.edu.cn

Abstract. At present, there is not a united, special law of Internet Service Providers, and there are many shortages in existing law of Internet Service Providers. For example, lacking of the subject system of Internet Service Providers; not forming the internet law system; lowering the legal grade, and so on. As the important status of Internet Service Providers in information spread, whether the rights and obligations are bright or the behavior regulations are perfect, is directly related to the transaction order and development of internet market. In order to regulate the behaviors of ISPs and promote the healthy and rapid development of internet economy, the classifications of Internet Service Providers are given and their tort liabilities are discussed in the paper.

Keywords: Internet Service Providers; internet technology; tort; liability.

1 Introduction

With the rapidly development of E-commerce in our country, the new commercial transaction mode greatly stimulates the earnest and desire of deal, and propels the flourishing and development of economy. In the meantime, tort and internet economy are intergrowth. Some unfair competition behaviors, such as issues on internet group-buying, online water army, online posting to slander the opponents, paying for deleting a message, attract the people's attention. Facing various kinds of tort, people gradually turn to Internet Service Providers (ISPs for short). It is a hot spot that how to enhance the regulation on ISPs and manage the tort.

In the Belgian case of Sabam V.Scarlet/Tiscali, the Court of Appeal of Brussels asked the European Court of Justice to issue the preliminary ruling on the relation between the various the European Directives, which together make up the legal framework regarding the liability of online intermediaries. Taking into account the many ambiguities that currently exist in this legal framework, the decision of the court of justice will be of great importance on the evaluation of the rights and obligations of ISPs [1]. The only way to resolve the liability of ISPs can be divided into two stages. The first one is whether self-regulation or co-regulation, the interests should be protected. The second one is to make the elected regulatory model effective

[*] Xiaomei Hong is a lecture, Dr. and teaches commercial law and economic law.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 475–480, 2011.

and each regulatory activity such as standard setting, monitoring and executing be considered separately. At this stage, it is important to bear in mind that private regulation, public regulation and tort law are complementary strategies and not to stop at the formalism of tort law rules [2]. The essence of eBay' service consists of mediating between sellers and buyers. According to the court, eBay has developed a profitable commercial activity in online auctioning and is primarily an auctioneer. Although the court agreed that eBay can be partially considered a "hosting provider" (Because it stores information regarding the items put up for sale), it determined that eBay's activities as a hosting provider and as an auctioneer can not be separated from one another. Since eBay merely stores information regarding the auctioned items in view of the auctioning process, eBay does not enjoy limited liability in its capacity as auctioneer, and thus falls under the general regime of liability for all of its activities [3]. Evidence shows that internet has become a more powerful medium than the traditional mass media. Interestingly, the research regarding the management of brand power by ISPs is scarce. Namely are brand image, brand awareness, brand trust, brand preference, and the relation among the customers' perceived value. Research turns out that the essence of branding is the choice of ISP services. Moreover, the relative impact of brand awareness and brand trust on brand preference is also consistent [4].

Some scholars divided ISPs into internet hardware service providers and internet soft service providers. The former contained the internet access providers and internet storage space service providers. The latter included internet content provider and internet technology provider. Because their behaviors are different, elements of infringing act are various [5]. Some scholars think that we should learn from foreign countries, divide the form of ISPs, namely internet access provider, search and links provider, internet storage space service providers, and apply different rules on different kinds of ISPs [6]. Some scholars divide ISPs into search engine, IM, Mobil internet services and web portals according to the main business [7]. Some think that ISP which provides technical support and communication platform can easily cause the tort of links, storage service, search, circumvention of technological measures [8]. Understanding the ISP liability from narrow angle, some scholars think it as omission tort liability, fault liability which violates security obligations. ISP undertakes the fictitious joint liabilities with the internet users to the victims [9]. Aiming at the different service nature of ISPs, some suggest that it is appropriate and necessary that ISPs supervise the clients and stipulate the different tort according to the nature of tort [10].

From the above mentioned literatures, the categories of ISP are not uniform and different. Mainly categorize the subject types from technology, litigation and service, and then propose the different tort. Generally, the tort of ISPs is limited to the fault liability. In view of the rapidly development of E-commerce in our country, the network violation is more and more serious, and it is urgent to regulate the transaction behavior and order. Based on the E-commerce transactions reality, the need of network economic development, and referencing the good legal idea and experience in foreign country, this paper defines ISPs at the commercial subject angle, clears their responsibilities, perfects the internet economic legislation, and ensures the safety and order of E-commerce transactions.

2 Insufficient Legislation of Internet Service Providers in Our Country

In 1998, our country finished the first internet transaction, since then, started our country's internet economic times. Prior to this, some terms involved ISP in laws, such as constitution, civil law, criminal law, rights law and so on. However, after all, ISPs are different from the traditional commercial subjects. With the appearance and development of E-commerce, our country successively formulates some systems which regulate ISPs, such as Administration of the Chinese Public Multimedia Communications, Internet Information Services, On the trial of cases involving Computer Networks for Copyright Disputes the Interpretation of Legal Issues, Information Network Transmission Right Protection Ordinance, Network Behavior of Commodity Trading and Related Services Interim Measures and Tort Laws, and so on. These laws provide guarantee for the healthy development of our E-commerce. But there are many shortages in legislation. Mainly for the following aspects:

2.1 Lacking of Internet Service Providers' Subject System

Our ISPs mentioned earlier legislation in 1997 was Administration of the Chinese Public Multimedia Communications, which regulated the internet access providers and information source providers. In 2006, Information Network Transmission Right Protection Ordinance was implemented, which stimulated ISPs who provided information storage space, search, links service, and their legal obligations and liability in the fourteenth item. In 2010, the tort law regulated the legal responsibilities of ISPs and internet users, and jointed liabilities between them in the thirty-sixth item. But it does not define and explain the subjects which ISPs contain, so that judicial determination is different to ISPs, and their liabilities are not the same. The reasons of this kind of situation are that our country pays more attention on the internet market order, pays less attention on the rights and obligations of ISPs, and it results in the unclear classification, and a vague definition. The fact has been very clear, but disputes continue.

2.2 The Internet Legal System Have Not Been Formulated

Legal system is a unified whole of all the existing laws according to the different legal department, which is consistent coordination between law act upper and lower, bridging coordination between branches of law, substantive law is comprehensive and detailed, procedure specification should ensure the rights to be finished. As the law to adjust the internet economy and market, it should form its law system. However, from the existing law, there are some issues, such as no coordination between law and law, shortages of substantive laws. We have not drawn up the internet basic law, for example: E-commerce Promotion Act. We have not enough ISPs system. What happens in the market, what legislation regulates. There is no forward-looking and overall planning in internet law, and then it results in the relative chaos in internet law. Internet law system is further strengthened.

2.3 The Lower Legislative Level

Legislative level standards for its authority, and shows a certain degree of stability. Nowadays majority of internet legislation in our country are management approach, ordinance, interim measures. In one hand, it indicates the lower legislative level, lacks in authority and stability. On the other hand, it illustrates that the recognition to new internet economy and young internet market needs to be further. We trust that our country will gradually promote the legislative level with the comprehensive understanding to the internet market.

3 Determining the Categories of Internet Service Providers and Tort Liabilities

The same responsibility principles with the international rules, our laws adopt the fault liability to ISPs. According to the civil law, whether the perpetrator has fault depends on whether he can foresee the consequence of his action. There are differences between the general predictable level and the professional predictable level. ISPs are the important information communicates media in internet space. Based on the development of internet technology, our country's current legislation and practice, and the foreign advanced network legal experiences, ISPs are divided into two categories, they are Internet access provider and host service provider.

3.1 Internet Access Provider

The Internet access provider provides the cables and switches for the information dissemination and gives the access services for Internet. Because it mainly offers the basic facilities for networks, and this is equivalent to the pipe of information dissemination, automatic transport service, it can not edit the information, and could not decide who is the information sender and receiver. It has no fault as the subjective, so it does not assume responsibility for infringement which the internet user violates others legal rights using the network basic facilities. But it should undertake the obligation to assist the infringer to stop the infringement actions in technology and facilities.

3.2 Host Service Provider

The host service provider is the provider of the network platform, includes the storage space, real-time communication, searching, links, upload and so on. They can edit some information that flows through their systems in technology, and have certain ability of control. So they should undertake the prior review obligation and the afterwards control obligation. The host service provider can be divided into the following categories:

(1) Information storage space service provider

It provides the service of browsing information, uploading information, and communicating information in real-time, such as BBS, chatting room, Email and so on. They undertake different liabilities under different circumstances. In Information Network Transmission Right Protection Ordinance, they do not assume prior review

obligations, but should take reasonable measures, such as deleting or disconnecting of infringing content to stop the dissemination of infringing content. If receives the notice of victim to delete or disconnect the infringing content, ISPs are failure to fulfil the legal obligations, then they should undertake joint liability with internet users who implement infringing behaviour. However, some information storage space, such as Baidu Library, in its "Top Recommended" part, whether the names and covers of some documents are the published famous works can be verified easily in general predictable level. With the professional knowledge and experience, Internet Service Providers can decide the works are not permitted uploaded, and delete the document immediately. Otherwise, they should undertake tort liability. It should undertake tort liability when they know well or should know that the dissemination of the listed document is not permitted. It is necessary to ask ISPs to assume prior review obligations, namely before informed existing infringing information, ISPs should initiatively review the legitimacy of the information in their system. It is feasible on technology, and can not increase the cost of information dissemination. Meanwhile, it is good to encourage and promote the development of internet economy.

(2) Searching and links service provider

Search engine classifies the information in network through the special process, helps the users to find the needing information, namely information search. Links is the service that directs from a web to a target. In Information Network Transmission Right Protection Ordinance, they do not assume prior review obligations, but should take reasonable measures obligations. If know well the linked content is infringed, they should undertake joint liability. Because of helping others fulfilling the infringing behaviour via the network, Yahoo website was decided to undertake the tort. Yahoo website only deleted the infringing links that the concrete addresses are provided by the plaintiff, idled in deleting the other infringing links, and didn't carry out the afterwards control obligations. So it should undertake the tort liability for its subjective fault. In fact, the situation is much more complex. For example, "it is quite different between the two characters of acts which provide 'empty search box' and 'list'. We don't only think the service as indirect infringement because it lists the effects according to the items by 'empty search box', but also decide the information whether plays a role in locating the infringement content, then determine whether service provider should disconnect the links. However, as to 'list', we should decide the subjective fault of ISPs according to the list names and content [11]." So, shallow link and deep link have the different legal significance.

4 Legislative Proposals

4.1 Formulating the E-Commerce Promote Acts

Referencing the developed country's experience, we should introduce the E-commerce Promote Acts to set up the basic law of E-commerce transaction and service. Usually, the basic law is to adjust the kind of social relationship, which regulates the general rules, mainly subjects of legal relationship, right and obligation and so on. The market needs to be ruled by the law, and it ensures the determinacy and foresees ability of market transaction. The existing network legal sources should be combed and integrated to build the internet law system.

4.2 Establishing the Internet Market Subject System

"Motive power of internet economic development is the open of internet terminal [12]." The information in Internet is voluminous, and it offers possibilities to the selection of the information. Meanwhile, the information is mixed so that lots of infringements exist in market. People know that it is a key to enhance the liability of ISP in order to regulate the network navy. So, it is important to clear ISP's type, right, obligation and liability. Judicial interpretation needs to clear the types of ISPs further in tort law to determine the liability subjects.

5 Conclusions

Internet Service Providers are one of the important commercial subjects in network market. Internet market transaction has been completed by many interconnected ISPs. Their services are different, and tort liabilities are various. We should clear the categories of ISPs, and their tort liabilities. It helps in solving the disputes to maintain transaction order and safety of the network market, and facilitate the network economy development. It is urgent to issue judicial interpretation to perfect internet legislation.

References

1. Patrick, V.E., Maarten, T.: Recent Events in EU Internet Law. Journal of Internet Law (July 2010)
2. Sophie, S.B.: International Review of Law. Computer & Technology 23(1-2), 153–165 (2009)
3. Patrick, V.E., Maarten, T.: Recent Events in EU Internet Law. Journal of Internet Law (October 2008)
4. Kevin, K.S.C.: Power of Branding On Internet Service Providers. Journal of Computer Information Systems (2010)
5. Lan, X.W., Peng, X.K.: Analysis on "Internet Service Providers". Technology and Law (2009)
6. Liu, Y., Huang, Q.: The Liabilities of Internet Service Providers in Tort Law. Journal of Jinan University(Philosophy & Social Science Edition) (3) (2010)
7. Wang, L.: A Research on Operation modes of Chinese ISPs. Modern Telecommunication Technology (5) (2007)
8. Chen, H.Z.: Analysis on the Tort Liabilities of Internet Service Providers. China Intellectual Property Report(2009)
9. Zhou, Y.Z.: Tort Liabilities of Internet Service Providers. Theory Research (3) (2010)
10. Liu, D.L.: The Appropriate Supervision Obligations of Internet Intermediary Service Providers—Brief Review the Tort Law[EB/OL], http://www.apcyber-law.com/
11. Wang, W.: Research on Information Service Provider—Comments on the Verdict of Fanya V. Baidu[EB/OL], http://www.fengxiaoqingip.com/ipteseluntan/luntan3/lt3zeren/zrzonglun/20090721/4787_9.html#
12. Fu,C., Zhang, G.Q. Yang, J.: Thinking of High Autonomy of Internet Architecture [EB/OL], http://ict.cas.cn/kxcb/kxr/201009/t20100907_2945906.html

Practical Teaching Problems and Countermeasures in the Major of Undergraduate Human Resource Management

Xiao-rong Jiang and Jin-liang Li

School of Business and Administration Management,
Xi'an University of Technology, P.R. China, 710054

Abstract. The major of Human resource management as a practically profession, has a late start in our country, but its development is rapid and its social demand becomes intense. The practical teaching in human resource management as an essential part of professional education in the major of management, its application determines the importance of the links of teaching practice in the whole teaching system .it has already been taken more and more seriously by various universities, students, employers and the relevant researchers. Aiming at the undergraduate practical teaching in the major of human resource management, the article analyzes the practical teaching process throughout existing problems in the school laboratory construction, the practice base construction, practical teaching system settings and the equipment of teachers' team. And also in the base of this, it proposes countermeasures and suggestions about practical teaching in the human resource management.

Keyword: Undergraduate, the major of human resource management, practical teaching.

1 Introduction

With China's economic development and intensified efforts in opening to outside world, the currently social demand in cultivating inter-disciplinary talents and applied talents puts forward an urgent requirement about higher education, while the traditional talent training mode which attaches more importance to theory but ignores the practice is facing severe challenges. But college students' practical teaching which is an effective way to getting rid of the higher educational dilemma enables the student to improve independent thinking, independent research and independent ability to explore questions and have access to professional skills and innovative ability on the basis of deepening understanding in theoretical knowledge by providing practical platform realistically at the same time [1]. In recent years, when the major of human resources management has developed very quickly, universities are thinking the same question: how to reveal their own characteristic to finish the task of cultivating applied talents in the education and teaching. At the same time, because of the particularity of the human resource management major, it makes the practical teaching for talent training is extraordinary. However, in the actual operation process, there still exist many questions about the practical teaching in undergraduate human resource management, which has

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 481–486, 2011.

impacted on the quality of talent training in the major of human resource management. This article will do the corresponding analysis and put forward the countermeasures to improving practical teaching in human resource management.

2 Analysis about Practical Teaching Problems in the Major of Undergraduate Human Resource Management

2.1 The Construction of Campus' Professional Laboratory Behind Professional Development

The development of practical teaching must have certain environment and the foundation, so the establishment of the professional laboratory can provide students with good learning environment and combine theory with practice better together, which is very beneficial to enhance students' ability to analyze and solve problems, and which is also good for improving their ability to make product sample by hand. However, that many domestic colleges which open the major of human resource management haven't established professional lab leads to the situation that students of the major have not seen personnel test and evaluation software and human resources management software which they just knew and heard [2]. But now units usually require students to manipulate expertly as they have taken their quarters, which usually have the specific embodiment in personnel application. Some domestic universities have seized opportunities and established HRM comprehensive skill training room, but that experimental teachers which are qualified, proficient in HRM is scarce and that curriculum design of related courses is not enough lead to students' wallowing dates whole, not fully understanding and failing to reach the effect in professional training class.

2.2 The Shortage of External Training Base

Human resource management is an operational profession, such as job analysis, staff recruitment, training management and performance management, etc. If the students don't have the opportunity to experience knowledge, inspect knowledge and skills by themselves, what they learnt are difficult to become their ability [3]. Therefore, establishing external training base are particularly important. However, due to various reasons, they failed to establish enough external training base. Even external training base has been established, most is in tatters and fails to play the role of practice base due to a lack of attention and management. How to make the training base function fully and effectively is a worthy of further exploration.

2.3 The Unreasonable Settings of Practice Teaching System

How to set up system of putting practical teaching into teaching plan which is the overall program of teaching is directly related to the practical teaching effect. Combining practical teaching with the theory teaching can ensure students to put theory into practice and deepen the recognition and understanding of the theory in practice [4]. But that many settings of undergraduate majoring in human resource management on the practical teaching is not reasonable shows as follows: unreasonable time

arrangements, the open of practical course before some basic and theoretical course, unreasonable practical teaching content, much theory but less practice and the low standard of practice even if a certain course needs the arrangement of course design. These problems will impact the practical teaching effect.

2.4 Professional Teachers' Operation Abilities Should Improve

As the rapid development of China's higher education in the last ten years, The main sources of college teachers in China are college graduates who may have master or doctor degree and have a rich theoretical knowledge but lack of work experience in the enterprise, never engaged in company, enterprise operation and management experience, so it will unavoidably lead to theory from practical tendency [5]. Therefore teachers guiding students in practical measures seems to be not enough "strong teaching" , making practical teaching effects discounted. In order to adapt to the needs of professional development, many teachers are added to the professional teachers' team who were in economics, specialty related to management [6]. However, for the major of human resource management, they belong to “switching to a new profession”, their practical teaching experience shows more deficiency. They are often accustomed to the traditional "spoon-feeding" teaching and ignore the training of students' ability to practice and their innovative ability, which also impact the practice teaching effect.

3 The Strategies and Recommendations about the Improvement of Human Resource Management Practice Teaching in the University

3.1 To Change The Education Philosophy

The reform of Human Resource Management practice teaching must change educational ideas and educational philosophy at first, to foster “thick foundation, wide caliber, high capacity” of high-quality, innovative compound talents as a guiding ideology, to the student as main, gradually realize the conversion that professional education to basic quality education, imparting knowledge-based to capacity-building, the common education to personnel education, the importance to systematic theory to avocations of a comprehensive application.

3.2 To Establish A Modern Human Resource Management Lab

To establish a modern Human Resource Management Lab, The first is to build a team which should be structured, skilled and could provide good protection for the reform of practice teaching and technology innovation. That not only can undertake the experimental technique work in the laboratory such as equipment public service, the development and maintenance of equipment, participation in the reform of practice teaching, providing important protection for the teaching, research and social service. But also can carry out practice teaching, experimental techniques or laboratory construction and management research, and integrate theory with practice, sum up

experience, write a research paper, or take part in the preparation of experimental materials and so on. At the same time, once the persons who will be responsible for the work in Human Resource Management Lab is designated, then job training or job training will be needed to enable them to have enough knowledge of Human Resource Management. These persons should understand to ask the software company to equip Human Resource Management teaching software and experiment instruction in the procurement of software, also know where the problem is and communicate in a timely manner, requiring appropriate and timely service during the period of maintaining the normal operation of software. Second, in the Human Resource Management lab, a complete and real copy data operations in all aspects of Human Resource Management should be equipped, such as corporate Human Resource Management department work plans, the statement of work of enterprises, labor contract, employee handbook, recruitment plan, pay management practice, evaluation and assessment program, employee training management approach and so on. Third, Content in the experiment, it should be focus on job analysis, comprehensive assessment of senior personnel, performance evaluation, salary system. The main experimental approach should be working simulation and making the laboratory a three-dimensional simulation-based professional company, so that it can realistically simulate the actual operation of the company as well as specific aspects of Human Resource Management, so as to truly cultivate Human Resource Management students practical skills.

3.3 Establish Stable and Long-Term Social Practice Base

The external practice of undergraduate human resource management is one of the important links of practical teaching. To make external practice better, we should choose the higher level of human resource management or representative enterprise in industry, applying the principle of the mutual benefit and common progress to establish stable and long-term social practice base, which is an effective method of strengthening school cooperation, developing human resource mutually and achieving mutual benefit and development at present. The school makes use of existing resources and conditions of the enterprise in order to establish a relatively stable internships and employment base and accomplish the teaching task better. In this process, it can inspect teaching achievement, get the advanced information of enterprise management timely, revise and adjust teaching goals timely and strengthen the cultivation of students' target by students' quality and ability performance in practice, to make schools form their own characteristics and advantages in the fierce competition in the market. Enterprise can make technology development with schools, gaining new needed technology, achievement and information of development. And at the same time, in the process of production it can evaluate students' comprehensive quality and ability to choose excellent employees suiting themselves, reducing production cost and employing risks, enhancing talent advantage and competitiveness. In addition,qualified schools can also open consulting companies, which own the function of practical teaching in the process of providing consulting service activities. In the practical teaching process the teachers should be equipped with .At the end of practice, teachers should ask students writing practice summarizes which is identified by practice units and guiding teachers.

3.4 Organize the Students to Carry Out Social Survey in a Planned Way in Their after School Time and Summer

Social practical activities which establish a complete set of assessment and incentive system to actively encourage students to attend these kind activities, for example ,"a program under which officials, doctors, scientist and college students go to the countryside to spread scientific and literacy knowledge and offer medical service to farmers" coordinating with schools and youth corps committee organization of college students, make the students free teams according to their own interests, free proposition, take full advantage of vacation time, go deep into society, increase cognition of China's national condition and society, accumulate social experience, do a new understanding about the importance of the course and also lay a foundation to learn the following curricula.

Besides, pay attention to the course interspersed with cognitive practice. According to the characteristics and schedule of each course, systematically organize the students to understand the condition and carry out enterprise market surveys by using spare time, for example, we can organize students to do on-the-spot investigation of the enterprise before each professional course or during the lecture in order to increase students' perceptual cognition. And also the experienced chief of the department of human resources can be invited to the classroom, conducting seminars, introducing the application of human resource management in the enterprise, etc. the above methods as an effective complement of class time can improve the learning interest of the students, increase the richness and interest ,and they also can make up the inadequate intuitive feel of laboratory practice.

3.5 Strengthen the Construction of Qualified Teachers and Improve the Ability to Teach Practically

The school should pay high attention to teachers' training of the practical teaching and echelon construction. On one hand, the particular way to using the policy guidance and attracting high level teachers engaging in practical teaching work is as follows: By the personnel department of the school coordinating, the school should organize teachers, especially the professional backbone of teachers, to work in enterprise as a corresponding position of enterprise management, which makes teachers increase more perceptual knowledge of human resource management, know the specific operation of it fairly well and combine the their theory knowledge with their practical knowledge effectively. On the other hand, encourage teachers to go in for the investigation of the enterprise' scientific research and participate in the enterprise' work of consulting and training to enhance teachers' practical ability and creative ability. In addition, Hire experienced minister of the human resources department and the staff of human resource management as the guiding teacher to supervise and control students' practical work. In order to make the guiding teachers have a sense of responsibility and mission, the school could give them letters of appointment and pay them corresponding compensation. The school can also employ them as part-time professors undertaking practical teaching task on campus.

Strengthening graduate practical teaching of human resources management not only can make students grasp the knowledge and skills of human resources management more deeply and more concretely but also can promote students' ability

to solve problems comprehensively, harmonize interpersonal relationship and work collaboratively ,offset the shortage of their less working experience and enhance their social flexibility and workplace competitiveness. Because of those students may be far more coincident for the talents standards, the training will be good preparation for their future job. Therefore, it is imperative to strengthen the practical teaching.

References

1. Zhu, X.: The Problems and Strategies of Human Resource Management Practice Teaching in the University. College Education (March 2007)
2. Liu, X.: The Organization and Implementation of Human Resource Management Social Practice Teaching. Population and Economy (October 2002)
3. Nan, L., Xiang, J.: Strengthening Human Resource Management Practice Teaching. Hei Longjiang University Of Law (May 2005)
4. Liu, F.: The Necessities and Uses of Strengthening Human Resource Management Practice Teaching. Higher Education Forum (March 2006)
5. Liao, Y.: The Research of Human Resource Management Practice Teaching in the University. Guangxi Institute of Finance (October 2006)
6. Liu, Y., Lin, Y.: The Discovering of Human Resource Management Practical Teaching in Local Colleges. Education in China (August 2007)

Investigation on the Training Mode of the Computer-Application Talents Based on the Information-Technology Orientation

Haicheng Xu, Jinwei Fu, Bo Yang, and Rui Li

Eengineering College, Honghe University,
Mengzi, Yunnan, 661100

Abstract. This paper analyzes the problems existed in the personnel training of the computer science major in the newly-built undergraduate college, according to the teaching reforms of the Computer Science and Technology major in Honghe College, investigate the training mode of the computer application talents based on the information-technology orientation, it specifically including the following aspects: the train of thought in the talents training mode, personnel training mode, the design of the talents training program, the concrete measures in the reform of the talents training mode and so on.

Keywords: Information-technology Orientation; talents training mode; application type.

1 Introduction

With the development of computer and communication technology, China began to enter the information-based society. Using informationization to stimulate industrialization and building a moderately prosperous society in all respects has become a basic national policy. According to statistics, there are 847 undergraduate colleges with computer major, and with an enrollment of nearly 40 million, it is the major with the largest scale all over the nation, and transfer lots of talents for the society. However, with the popularity of computer technology, "be able to use computer" is no longer the unique advantages for the computer science majors. In addition, the connotation and extension of the "Computer Science and Technology" which was formed a decade ago has a great distance between the anticipation of the society to the “computer talents” which was caused by the highly development of the information and technology. At present, the job-hunting for the computer major is becoming more and more difficult, but the need for the computer "application-type" talents is urgent in all walks of life. Thus, most non-key institutions, especially to those newly-built undergraduate institutions, they must create personnel training programs with features based on their own status and advantages. In this paper, I take the talents training for the Computer Science and Technology major in Honghe College as example to explore the train of thought for computer application-type talents.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 487–492, 2011.

2 The Existing Problems of the Talents Training of the Computer Major in Newly-Built Undergraduate Institutions

2.1 The Profession Orientation Is Out of the Reach of the Social Need

The computer science major in colleges of all types and of all levels all carried on Specialty Setting based on the undergraduate specialty catalogue in 1998, the biases exist between professional orientation and social demands, which is mainly showed at the following aspects: using single professional setting to deal with a wide range of subject areas, and using single training standard to deal with a variety of social needs. In view of the knowledge structure of the computer science major, it involves hardware, software, network, communication, multimedia and many other disciplines, to cultivate a "all-rounder" in the computer science major according to one profession in prescribed educational system only is obviously impossible. In view of Computer applications, it relates to industry, agriculture, business, transportation, government, research institutes, schools and many other enterprises, administrative institutions and fields. The attempt to use a single training standard to meet so much social needs is obviously inappropriate.

2.2 The Growing Pressure of Employment

With the expansion of college enrollment, a large number of computer graduates flock to social, then the employment situation is more and more severe. On the one hand, it is difficult for the computer science graduates to find a job, on the other hand, the enterprise could not find the talent to solve practical problems, which is caused by the fact that the graduates possess a comprehensive knowledge structure but can not operate or implement specific projects in work.

2.3 Teaching Content Revolved in the Curricular System Is Lagged Behind

The teaching for the computer science major in newly-built undergraduate colleges, because of the lack of the development track for new technologies, obvious blindness of the design of the teaching contents existed, the teaching plans is similar to the old and famous undergraduate colleges, although the curriculum system setting is comprehensive, it has no feature at all.

2.4 The Construction of Teachers Quality Is Lagged Behind

The teachers teaching at the first line in the newly-built undergraduate school are mostly young teachers, their teaching task is arduous, and have no chance to develop and enhance themselves further, it is difficult for them to have the energy to complete the task of teaching creatively. As the rapid development of IT industry, knowledge and technology is changing rapidly, which caused great technical pressure to professional teachers.

2.5 Weak Practical Teaching Procedures

Various experimental apparatus and equipment in newly-built undergraduate colleges and universities can not meet the needs of practical teaching after the expanding

enrollment at large scale, the experimental procedure must be enhanced, and high-quality experimental design and the configuration of qualified experimental guiding staff did not keep up with the demand. The practical ability of most students are weak, they lack systematic and comprehensive practical training, which caused serious problems in their practical operation skills.

3 The Construction of the Training Mode of the Computer Application Talents Based on the Information-Technology Orientation

3.1 The Design of the Talents Training Project

The Positioning of the Talents Training Mode

The training of the computer application talents based on the information-technology orientation is mainly taking the training of the computer information system integration technology as the main objective, to meet the needs of informational construction in a variety of enterprises; the training of the software development ability for students is only aimed at small systems, for large and medium-sized system, they only requires student be able to use and maintain.

Formulate Talents Training Goal

Train the application type talents that mastered the natural sciences and mathematics knowledge related to information technology, the basic theory of the computer science and the application knowledge of hardware and software, and that have the potential to apply this knowledge to information system construction and integration creatively, with the ability to understand, evaluate, select and implement, and be able to adapt to the changing needs of the society's technological progress.

Set up Professional Curriculum Project

According to the requirements of the profession norms set by *Higher Education Law* and the Computer Science and Technology Major Teaching Steering Committee of the Ministry of Education, to determine the student's knowledge, ability and quality structure. Implementing the training mode of "common courses, academic platform courses and professional courses are equal ", and developing a corresponding system of professional curriculum, and controlling the total credits of different courses within 160 credits, which including: common courses (56 credits) account for 35% of the total 160 credits, academic platform courses (22 credits): account for 13.75% of total 160 credits, professional courses (82 credits): account for 51.25% of the total 160 credits. Among the three major platforms including practical teaching courses (altogether 64 credits): account for 40% of the total 160 credits.

The Construction of the Practical Teaching System

In accordance with the relevant requirements of *The Development Strategy Research Report and Profession Norms (Trial) of Computer Science and Technology in Higher Education Institutions* and *Common Core Knowledge System and Curriculum of the Computer Science and Technology Profession in Higher Education Institutions* built a practice teaching system of the following four levels: “curriculum

experiment - curriculum design - practical training – graduation design", and have specific requirements to the curriculum experiment, curriculum design, practice, graduate design involved in the professional curriculum.

3.2 Train of Thought and Specific Measures in the Training Mode Reform

Construct the talents training system of "student-centered"

Establish the concept of "student-centered", adjust talents training projects and construct curriculum system according to the concept of "quality education" + "professional education", promote the talents training mode of " student-centered" training model, establish a teaching management system and evaluation system of "student-centered".

Emphasizing the dominant position of students, emphasizing the independent learning of students, training the lifelong learning ability, and the autonomic learning of students will be the main form in teaching procedures, teachers' teaching activities are mainly have guidance on the student's independent learning and personal development.

Learn new professional norms, lay foundation for the formulation of the talents training projects

The new professional norms was unveiled in 2006, not only people outside the field know little about it, but also the computer science teachers know little about it. Therefore, our school organized teachers in this profession to learn it, and research the obligatory knowledge field, knowledge unit, knowledge points, so to make most of teachers recognize that to choose the information-technology orientation to set the talents training projects in the newly-built undergraduate colleges is in line with the requirements of the applied talents training.

Emphasize engineering applications, strengthen teaching practice

The application must be emphasized, for the purpose of cultivating application type engineers. In each course, arrange a small curriculum design, and for those courses that have strong practicality set a large-scale curriculum design, so as to train students' integrated application capabilities fully. Ensure the corresponding practical conditions so that every student can be trained strictly in engineering practice. Here adding two ways of practice: 1) to practice outside the province; 2) long-term practice inside school. Graduation design mainly take the comprehensive training for the following: 1) topic-choosing; 2) surveying; 3) familiar with the problems; 4) search the information; 5) needs-analyzing; 6) set a research plan 7) summary design; 8) detailed design; 9) specific implementation and debugging ; 10) document writing; 11) issues and plans in written and oral expression ; 12) thesis writing. Try to train student's ability to research and implement, ability to write thesis and express themselves orally.

Strengthen the Construction of teachers rank

Building a high professional title, highly-educated, dual-qualified teachers rank, by taking the advantage of the good relations with and industries to arrange some teachers to have practice and training in enterprises, try to train more "Dual-qualified teachers" with working background as engineer. Hire or the introduce high-level talents who have rich experience of engineering practice to the ranks of the professional teachers from domestic universities, research institutions, and the IT companies, and what's

more, hire part-time teachers in the field to participate in teaching and teaching reform, to promote the construction of the "Dual-qualified " teachers.

Reform the Teaching Methods and Means

By taking the amendment of the talents training projects as the entry point, by training the computer application-type talents in various fields of as the goal, by enhancing the application capacity and quality of student as the starting point, by enhancing the training of the practical ability of students as the focus, to carry on the reform of teaching contents, methods and means. In the teaching process, try to focus on the students learning ability, practical ability, and the ability to analyze and solve problems. Change the teaching methods of cramming, try to use multi-modes teaching methods, combine the teaching mode, practice mode, exploration mode, self-study mode organically , positively take the teaching method of heuristic, discussion, seminar, collaborative and case study, try to provide students with self-learning and practice space, and to expand their horizons.

Curriculum Teaching Reform

Integrating corresponding course modules, complete the school education in the form of "Theory teaching+ Experiment +school practical training", complete industry education in the form of "practice teaching + off-campus practical training ", and fully reflect the principles of mainly training the application talents. Promptly remove the outdated course content, increase new knowledge and new technologies of the computer industry, and enable students to understand the new developments and new requirements in the IT industry. Combine with the development status of the profession, organize academic echelon to plan meticulously, compile professional textbooks and teaching aids material. Since March, 2008, we have changed nearly 40 professional curricula in accordance with the requirements of the information-technology orientation, and according to the needs of the teaching practice divided the teaching syllabus into two parts, the theory and experiment parts.

Building practice and training base inside and outside campus

Established a relatively complete network experiment environment, deploy some basic equipments such as the servers, switches, routers and so on, provide networking, testing, network management and other related tool software to provide students a network operating environment. Establish a creation laboratory, provide a range of devices, let the students themselves design and program to realize the corresponding functions. Establish a comprehensive laboratory, to realize the virtual experiments of multi-courses. Establish an integrated information engineering practice training center in accordance with the simulation actual engineer and the requirements of the practical environment, which served as a place for students to design comprehensively, do practice in large scale, and have practical training. Cooperate with Beijing, Shanghai and Chengdu software training bases, let students practice outside province. Within the range of the Honghe State to cooperate with other industries (non-computer) and establish practice training base together, which can not only train the practical ability of students, but also can serve for the local economic development.

Innovation-led, and take the application project research as the main line

Stimulate students to think creatively, implement research-based learning, and select excellent students to do innovation activities, and apply the innovation

activities to actual use specifically. The students can select their interesting research projects according to their own characteristics, and research it as the Students Innovation Fund, they can also choose the research project from parts of the research topics of colleges, try to take the research of the actual application project as the main line, enable the students to master their professional knowledge comprehensively and systematically, and apply it into practice.

Combined with the local economic development, train application-type talents

For the purpose of training the application-oriented talents that can meet the needs of industries and companies, take the educational resources and environmental advantages of our college and the Honghe State industries, research institutions in the perspective of talents training, combine the process of talents training with the production environment scene organically. Make the development of computer technology available to students in all dimensions, and guide them to study with their own characteristics.

4 Conclusion

Honghe College is a newly-built undergraduate college of teaching type, which is mainly to train application-type undergraduate talents. The Computer Science and Technology major have explored on the training mode of the "application-type" since 2007, and now formed an application-oriented training mode of "general + platform + profession", the employment rate of this major have greatly improved in 2010. However, due to the new professional norms was introduced in 2006, many colleges and universities are still on the way of exploration, there is not much experience to refer; in addition, there is no large training base for the training of the application type computer talents in Yunnan Province, thus how to develop the practical ability of students under the current conditions is also a problem that needed to explore by ourselves.

References

1. Ministry of Education Higher Education Computer Science and Technology Teaching Steering Committee. The Development Strategy Research Report and Profession Norms (Trial) of Computer Science and Technology in Higher Education Institutions. Higher Education Press, Beijing (2006)
2. Ministry of Education Computer Science and Technology Profession Education Steering Committee. Common Core Knowledge System and Curriculum of the Computer Science and Technology Profession in Higher Education Institutions. Tsinghua University Press, Beijing (2007)
3. Chenming: Research on the Application-Type Talents Training of Computer Science and Technology Profession. Computer Education (6), 55–60 (2009)
4. Liao, X., Liao, Y.: On Exploration of the Application-type Talents Training Mode of Computer Science and Technology Profession. Computer Education (3), 32–34 (2009)
5. Mang, J.: Research on “3+1”Computer Profession Talent Training Mode In the Newly-built Undergraduate Institutions. Education Exploration (9), 31–32 (2007)

The Development of Digital Teaching Materials in Vocational Education —Using Electronic Engineering as an Example

Tien-Chi Huang[1] and Ping-Tsung Wang[2]

[1] Department of Information Management, National Taichung Institute of Technology, No.129, Sec. 3, Sanmin Rd., Taichung City 404, Taiwan (R.O.C.)
tchuang@ntit.edu.tw
[2] Department of Electrical Engineering, Kun Shan University
No.949, Dawan Rd., Yongkang Dist., Tainan City 710, Taiwan (R.O.C.)
wping@mail.ksu.edu.tw

Abstract. E-learning has broken the limits of time and space, brought a new face to teaching, and made learning more efficient and convenient. Many countries have taken the advantage of this "E-learning Fever" to promote their education levels. A great deal of schools, training centers, and cram schools have also actively participated in E-learning. They set up private platforms, and design digital teaching materials and courses, so that all sorts of digital materials can be found on market. However, most materials are noted for their quantity instead of their quality and are not designed under the regulations set up by a credible organization. As a result, some digital material designers are blamed for not improving the learners learning but wasting their time. The objective of the paper is to design a qualified digital teaching material which follows the certification regulation for digital teaching materials published by the Ministry of Education in Taiwan and is combined with our vocational engineering education, especially focusing on electronics.

Keywords: vocational education, engineering education, digital materials certification.

1 Introduction

The development of computers and internet technology has aroused the concept of "information integrated teaching" in the field of education, influencing not only the notion that students should be educated using computers and information technology, but also that information should be considered a learning tool. Current digital courses usually focus on actual lectures, with digital facilities as a supplement. Thus, regarding long-term objectives of digital education in its current stage is less likely to generate teaching interaction that breaks through the time and space barrier.

Due to these reasons, the Certificate Center for Digital Learning was established in 2009 to encourage all universities and colleges in Taiwan to offer online digital courses with a focus on quality and digital teaching materials to pursue the universalization, standardization, and excellence of higher education through digital education.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 493–500, 2011.

Numerous universities are currently actively promoting an asynchronous digital teaching platform, which assists teachers to incorporate digital education into their own teaching methods, and thus, is a useful teaching tool. However, under current circumstances, most teachers focus on mixed digital education, using an asynchronous digital teaching platform to assist actual courses. For example, platforms are used as a channel to give notifications regarding courses, to submit assignments and to share teaching materials. In conclusion, this study combined engineering education in vocational institutes, and followed the "Certification Regulation for Digital Teaching Material" to create online asynchronous digital teaching materials.

2 Research Background

2.1 Construction of Digital Teaching Materials

The Massachusetts Institute of Technology (MIT) planed to provide free access to all undergraduate and graduate courses before the end of 2007(MIT's OpenCourseWare). Online teaching then became a research focus, whose subjects have included the effectiveness of internet-based teaching [1], the construction of digital platforms [2][6], the design of digital materials [3], course development for digital learning, and the teacher's role in an online learning environment. However, previous studies on the construction of a digital learning platform have focused on the development of partial functions, without applying them to the overall process of digital learning. To design a comprehensive digital learning system, the platforms should consider both the teacher's and student's outcomes, as well as considering diversity and adaptability of materials, to improve learning effectiveness [4].

2.2 Digital Materials and Course Certificate

The development of digital courses in higher vocational education [5] has been highly valued during the transformation of vocational education. To encourage the sharing of resources, and to reduce learning gaps influenced by time and space, many higher vocational institutes have established internet colleges. These phenomenon imply the essential necessity of digital learning to higher vocational education. This study therefore designed new content for digital teaching materials by drawing from the experiences of teachers who have either obtained the certificate for digital teaching materials or are expecting to acquire the certificate issued by the Ministry of Education in 2011. We also provide several concrete suggestions for the certification standard and present implementation methods for the subject of electronics in engineering education.

3 Development of Digital Materials

3.1 Preparation for Certificate for Digital Teaching Materials

To reduce costs for system development and shorten development duration of materials, this study adopted the digital learning platform of an internet college run by

a university of technology in Southern Taiwan: "Sunnet Technology's Wisdom Master WM Pro 2" is a platform for certified digital materials and digital courses. Because the certification of digital teaching materials is complicated, this study required organizing standardized procedure of examination of digital materials (Fig. 1), developing digital teaching materials (first-stage), and proposing several concrete recommendations. The development plan of digital materials and various functions of the platform Wisdom Master are described in the following sections.

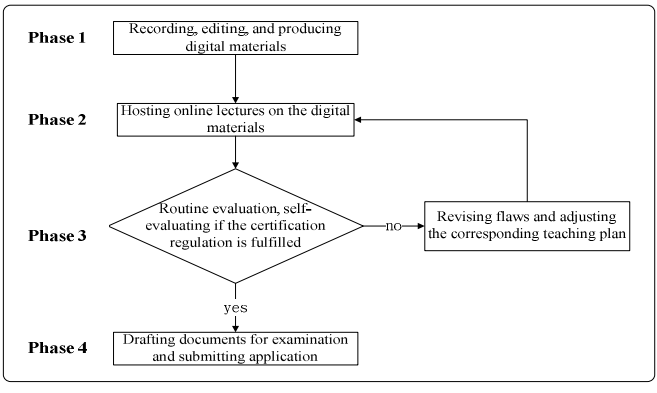


Fig. 1. Examination Procedure of Digital Materials

Firstly, actual materials should be digitalized, using "*.wmv" as the format of short films for class, "*.pdf" and "*.ppt" as formats for PowerPoint presentations in lectures, and other relevant digital materials used for a digital teaching platform.

In addition to an actual face-to-face lecture in the first week and an actual examination in the 18th week, lectures are all conducted online with synchronous, asynchronous, and mixed methods through a digital teaching platform. A routine assessment should be performed after a lecture in the 6th week and a questionnaire survey in the 18th week. Minor adjustments and revisions should be made to the digital materials according to the routine assessment (pretest) conducted in the 6th week. The questionnaire survey also serves as a basis for material editing and quality improvement. When the materials are thoroughly prepared and pass an evaluation, examination documents are drafted and applications submitted.

3.2 Self-evaluation on Digital Material Development

This study presents the implementation details regarding each regulation for digital materials announced by the Ministry of Education, selecting electronics (1) as the subject for the application of digital materials and courses, which is offered to sophomores in the four-year program of the Department of Electronic Engineering. The digital materials developed in this study should be taught with a combination of actual lectures and asynchronous distance lectures on a digital learning platform. The structure, file format, and procedure of each chapter of the digital materials are presented in the following figure:

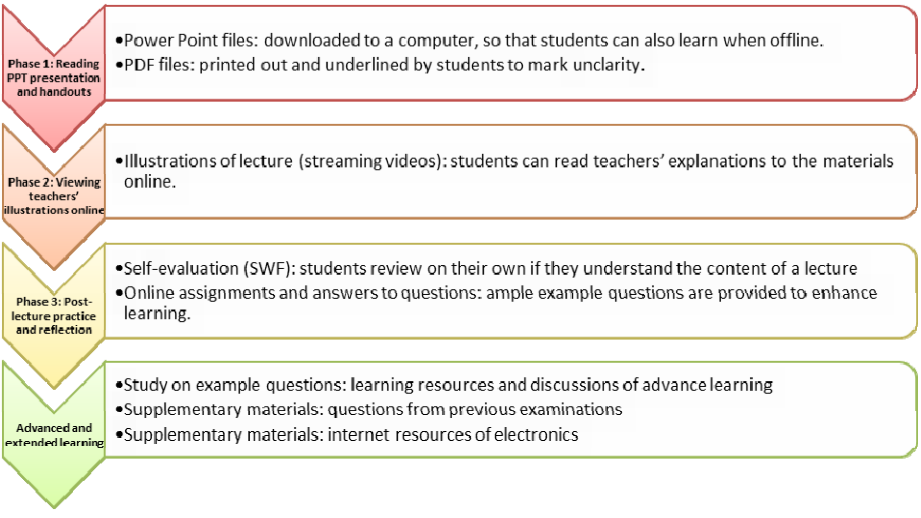


Fig. 2. Implementation Procedure of Certified Digital Materials

In 2011, the Ministry of Education drafted the certification regulation for digital teaching materials. Table 1 shows the division into four categories:

Table 1. Ministry of Education' Certification Regulation for Digital Teaching Materials in 2011

Certification Regulation for Digital Teaching Materials
1. Material content and structure (7 mandatory items and 3 alternative items)
2. Teaching design (6 mandatory items and 3 alternative items)
3. Supplementary design (1 mandatory item and 3 alternative items)
4. Media and interface design (5 mandatory items and 1 alternative item)

Four categories comprise several control items, mandatory and alternative. According to the certification standard published by the Ministry of Education, students must obtain an A (or above) on all mandatory items to pass the standard. An average result of A (or above) is required for all regulatory indicators (both mandatory and alternative items). In addition, an A+ on a mandatory indicator can counteract a B on an alternative indicator. This study used the A+ indicator of each control item as the objective of digital material development, and designed the materials as described below:

- **Regulation 1: material content and structure**

  Mandatory indicators: materials should demonstrate clear themes, unit names, numbers of learning hours, and applicable subjects, listing the knowledge, skills, and attitudes students should acquire. Material content should be correct, with appropriate proportions, a reasonable and clear structure, and order. They should include all teaching objectives, and conform to student levels.

To satisfy the requirements of mandatory indicators, themes and unit names should appear on the front page of the materials, and the number of learning hours should be included in course introductions. All materials are designed according to the mandatory course requirements of a department of electronic engineering in a university. Without confusion regarding structure and order, lecturers should introduce theories to students using simple language when teaching, starting from basic ideas to theoretical applications, to offer students with correct and appropriate content.

Alternative indicators: materials should clearly state the copyright of cited information, include latest content, and provide relevant learning resources.

In "course syllabus," writers and editors of the materials, copyright owners, and the information that the materials are available for use on this digital learning platform should be clearly stated to eliminate copyright disputes. The writers and editors should be responsible for updating the materials, to present the latest information on the course.

- **Regulation 2: material design**

Mandatory indicators: the design of materials should trigger the motivation to learn, and provide clear instructions for learning activities, examples, exercises and evaluations, and feedback of the learning activities.

To induce motivation, materials and lectures should be designed according to a progressive approach and the desire to explain complex theories using simple language. When studying course units regarding electronics, students should be provided with appropriate examples to learn from and practice diverse assignments, such as answering questions, design, and calculation.

Alternative indicators: in addition to listing content and recommendations for each stage of learning, materials should also provide appropriate remedial courses, and a supplementary FAQ.

The materials may include a course introduction at the beginning of each chapter, to introduce content, learning objectives, and the types of lectures of the course. The focus of the lecture and unit content should be summarized in the beginning of each unit, enabling students to grasp the major points, improving their learning efficiency.

- **Regulation 3: supplementary design**

Mandatory indicator: materials should provide a detailed content guideline.

Because the successful application of digital materials utilizes a digital learning platform, the two different types of users (teachers and students) require different versions of an environment handbook. A "common questions" section should be included, which organizes situations users often encounter, providing a detailed operational procedure and demonstration of questions concerning the system and materials.

Alternative indicators: materials should include a feedback mechanism, clearly demonstrating knowledge of the previously learned sections.

A digital learning platform should provide lecturers with individual e-mail addresses and discussion boards for lectures, where students can propose suggestions and pose questions regarding the materials. A learning platform can also

include a record of login details, as well as the time spent on a certain reading unit, to document the course names that students have already studied. Students can list the learned parts by browsing through previous courses.

- **Regulation 4: media and interface design**

Mandatory indicators: media used for materials should meet a certain standard, helping students comprehend the content through an appropriate display design, a consistent and convenient operation interface, and an applicable browsing tool.

Video materials are recorded using the software Powercam, which processes quality standards and lectures.

Only three pieces of software of Microsoft Media Player, Adobe Acrobat Reader, and Power Point are necessary for viewing multimedia files. Because operating and controlling the software is consistent and easy, students can decide the operational method for online learning.

Alternative indicator: materials should offer access to downloads.

The digital materials, including films, PowerPoint slides, and portable documents should provide easy access to downloads, attaining ubiquitous learning by enabling students to use previously downloaded materials when they cannot access the online learning platform.

4 Implementation Issues of Digital Teaching Materials

This study designed the teaching materials based on the aforementioned control items, and combined them with the digital learning platform "Sunnet Technology's Wisdom Master WM Pro 2". The actual implementation is described below:

- **Regulation 1: material content and structure**

To fulfill requirements, the digital platform introduces the learning objectives, applicable subjects, and an expected syllabus. In addition, a "studying guide" is provided before each chapter, to describe themes, learning objectives, and a recommended number of hours for learning.

The materials have three teaching objectives regarding electronics.

(1) understanding the basic structure and electric characteristics of a diode and a bipolar junction transistor (BJT).
(2) being familiar with the design technology of applied circuits of a diode and a BJT.
(3) being able to apply a diode and a BJT to design a practical analog circuit.

Learning units are designed according to the learning objectives. This subject is a mandatory course for undergraduate students in departments of electronic engineering, equipped with basic math and electrical circuits, ascertaining that the materials are in accordance with student levels. The materials are designed for a three-credit course, "Electronics(1)", to equip students with the basic abilities regarding the subject. Regarding lecture content, basic electrical circuit theories are reviewed in the first week. The physical characteristics of semiconductors that constitute electrical components are discussed in the second week, and a P-N junction

diode and its related applied circuits are analyzed between the third and the sixth week. Finally, from the seventh to the seventeenth week, the characteristics of a BJY, DC bias circuits, and AC small-signal analyses are discussed.

- **Regulation 2: material design**

To inspire students' motivation, the proposed design includes recorded streaming videos as multimedia materials, and self-evaluating tests. Besides, teaching schedule is listed in the "course syllabus." Second, to enhance student understanding concerning relevant topics to the course, the materials present appropriate real examples to be used as references in most units, having organized two practical assignments with a diverse format after lectures. Lecturers and the system will guide students to conceive solutions, in which students are given an opportunity for reflection.

- **Regulation 3: supplementary design**

To present content guideline, discussion boards, search function for internet-based courses, and title keywords searching functions are included in this proposed design. Moreover, to fulfill the indicator of the feedback mechanism, the platform also includes official and unofficial mechanisms. "Questionnaire function" serves as official mechanism, whereas unofficial mechanism includes Recommendations for the system and service center. "Recommendations for the system" allows both students and teachers can submit various suggestions and opinions. The "service center" provides the multiple contact information of platform managers.

To present detailed feedback mechanism in showing learned materials, the "my learning history" was designed to help with the identification of previously learned content, which contains last lecture, number of lectures, number of posts, number of discussions, and "reading hours". "Number of reading hours" shows which chapters (sections) students have already read and the number of hours spent reading the selected chapter. "Unread article" indicates materials, announcements, and discussions that have yet to be read, and calculates the amount of unread information.

- **Regulation 4: media and interface design**

The digital learning platform is categorized into two major sections: "personal information" and "campus." When students select an internet-based course to attend, the platform system displays three other sections: "interactive learning," "evaluation," and "information." The five main sections include corresponding secondary functions for student selection. These functions are precise and consistent, and therefore, easy to identify and operate. The operation interface and display design are consistent with the digital learning platform, with the left side displaying lecture units for selection, and the right side showing lecture content. Besides, this study produced slides of the teaching materials using the PowerPoint format. Moreover, high quality streaming videos are recorded using Powercam software. During the recording process, the materials are presented in a dynamic style, not only clearly delivering the content, but also enhancing student understanding.

5 Conclusions

This study introduced a drafted and edited design of digital materials based on the highest standard of the Certification Regulation for Digital Teaching Materials issued

by the Ministry of Education, examining the implementation of the 29 mandatory and alternative indicators of the four categories of the regulation, and provided an analysis. This study adopted the digital learning platform "Sunnet Technology's Wisdom Master WM Pro 2" for certified digital materials and digital courses, creating digital materials for "Electronics (1)" in vocational education. The production of digital materials can provide a sound digital learning environment that encourages students to learn independently and improve their learning outcomes. Future research will involve conducting a statistical analysis on surveyed data and an evaluation on learning efficacy, to boost the rate of success for students applying for the certificate. The proposed design and practical analysis of digital teaching materials can provide a complete understanding of digital teaching materials to researchers applying for the certificate, improving the quality of digital engineering education.

Acknowledgement

We wish to thank the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC 99-2511-S-025-002-MY2.

References

1. Tseng, K.H., Chiang, F.K., Hsu, W.H.: Interactive Processes and Learning Attitudes in a Web-Based Problem-Based Learning (PBL) Platform. Computer in Human Behavior 24(3), 940–955 (2007)
2. Kuo, Y.H., Huang, Y.M.: MEAT: An Authoring Tool for Generating Adaptable Learning Resources. Educational Technology & Society 12(2), 51–68 (2009)
3. Huang, Y.M., Chen, J.N., Huang, T.C., Jeng, Y.L., Kuo, Y.H.: Standardized Course Generation Process using Dynamic Fuzzy Petri Nets. Expert Systems with Applications 34(1), 72–86 (2008)
4. Chen, N.S.: Producing accessibility on using courseware on Web. In: 8th International Conference on Information Management, Taipei, pp. 667–674 (1997)
5. Yang, Y.C., Chuang, C.P.: The strategy of dealing with information technology integrated instruction from vocational high school. In: Faculty of Education of CUHK 40th Anniversary International Conference. Hong Kong Chinese University (2005)
6. Jeng, Y.-L., Huang, Y.-M., Kuo, Y.-H., Chen, J.-N., Chu, W.C.: ANTS: Agent-Based Navigational Training System. In: Lau, R., Li, Q., Cheung, R., Liu, W. (eds.) ICWL 2005. LNCS, vol. 3583, pp. 320–325. Springer, Heidelberg (2005)

3D Stratum Modeling Based on Profile Reconstruction*

Nian Yu[1,2] and Ning Zhao[3]

[1] Key Laboratory of Earth Exploration and Information Technology of MOE, Chengdu University of Technology, Chengdu, China
[2] China railway Eryuan Engineering Group CO.LTD, Chengdu, China
[3] College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China

Abstract. In the fields of petroleum exploration and engineering geology, complex 3D stratum model is always necessary to be constructed and used for analogue calculation or geological study. On the basis of analyzing features of 3D stratum modeling, this paper proposes a set of complete 3D stratum modeling algorithm based on the profile reconstruction. The algorithm can help the user to carry out interactive modeling by using the existing data simply with a high efficiency, and to construct complex model having fault; therefore, it has good practicability.

Keywords: 2D modeling, 3D modeling, block track, Delaunay Triangulation.

1 Introduction

3D geological model is generally complex, change of the stratum surface is big and there may also be surface intersection and offsetting conditions. However, the actual data for modeling is always limited; generally, only some 2D profile data or even a little bit of drilling data is available. The result of direct automatic modeling by computer is generally not satisfactory, therefore, the user needs to carry out interactive modeling relying on the own experience.

3D geological model is very different from the CAD modeling, and the biggest difference is that when carrying out geological modeling, the user has no complete and precise understanding of 3D model; 3D geological model is not to design stratum by the imagination thereof, but to represent the stratum model which can truly reflect underground structure with the original data. Therefore, it is inadvisable to design model in 3D view according to imagination or experience of the user directly. The author proposes a method of automatically constructing 3D geological model by using 2D profile, which enables the user to make full use of the existing data to carry out interactive design and generate complex 3D model having fault automatically. The algorithm needs the user to: at first, interactively edit a group of parallel 2D stratum profiles with 2D profile editor in a certain direction; then, generate initial model with reconstruction algorithm; at last, refine, cut and sew each stratum surface so as to generate smooth rational stratum model.

* Supported by National high-tech research development plan (863 Project No.2009AA06Z108).

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 501–507, 2011.

2 Profile-Based 3D Geological Modeling Method

In the railway investigation and construction process, because the geological object always locates underground, people have to obtain part of feature information of the geological object by means of drilling and geophysical exploration and so on, and to obtain basic feature of the object by analyzing, deducing and explaining the information. The perceptual process of people to cognize the objective things is always from 2D to 3D, and from local part to the whole, so the user always carries out analysis, explanation and deduction on the 2D profile at first.

2D profile edition is the first step of modeling, and the user needs to draw multiple parallel 2D profiles according to a certain direction. The specific profile edition operation always requires the user to draw stratum profiles according to original data and the experience, which may take the user much time, but this step is very important, because the accuracy of the 2D profiles judges the accuracy of the whole 3D model. At present, the 2D profile edition technology is mature, stratum curve is always represented by polygonal lines or cubic spline curve, which allows the user to represent various stratums and fault situation freely.

2.1 The Definition of 3D Geological Modeling

In the 3D geological space, stratum is always divided into a limited number of geological structure layer blocks by the fault, each layer block is an independent geological block, and each independent geological block has its own geological attribute and outer surface. Spatial 3D geological block is described by surface model, thus the 3D geological block is defined as: $g = H_u \cup H_d \cup S_b$, in which H_u is the upper interface of the layer block, H_d is the lower interface of the layer block, S_b is a closed boundary surface which forms the upper and lower interfaces of the layer block.

2.2 The Implement Structure of 3D Geological Modelling

Compared with 2D geological model, spatial distribution of earth surface, stratum and fault plane and combination relationships of the 3D geological model are more complex and more abstract, and it is more difficult to understand the topological relation. This chapter defines the 3D geological block on the basis of 2D complex geological profile modeling, describes the 3D geological block model using spot, triangulation, surface patch and 3D geological block. The 2D geological profile is composed of multiple 2D geological blocks which are sections of 3D geological blocks formed on the 2D profile; 2D geological blocks in multiple spaces can describe roughly topological relation of basic geometric shapes of the 3D geological blocks with the other geological blocks.

3 Profile-Based 3D Geological Modelling Method

Realizing 3D model reconstruction between two adjacent 2D geological blocks needs to use a series of triangular patches which are connected with each other to connect

the adjacent two outlines in space. However, we need to seriously study how to guarantee that the connected 3D surface model is rational and has good property. The numerous basic triangular facets formed by the control points on the outlines connecting the adjacent two 2D geological blocks should form 3D surfaces that are connected together and should not intersect with each other in the triangular faces. Therefore, the triangular face set is rational only if it meets the following two conditions:

(1) Each outline segment must appear in only one basic triangular face. Therefore, if the adjacent two outlines each have M and N outline segments, then the rational 3D surface model will contain M+N basic triangular faces.

(2) If one span is a left span in a certain basic triangular facet, then the span is and only is a right span of a basic triangular face. The triangular face meeting the above conditions can be called acceptable body surface. But it is obvious that, for the adjacent two outlines and point range thereon, there are multiple different combinations of acceptable body surfaces that meet the above conditions. How to judge one necessary combination among the multiple combinations of acceptable body surfaces, exhaustive search method is obviously not advisable; using synchronous progress method for reconstruction of outlines of the 3D geological blocks always brings a better effect, which can guarantee that connection of triangles can be carried out synchronously roughly and can also guarantee a good connection for stratum outlines with bigger offsetting.

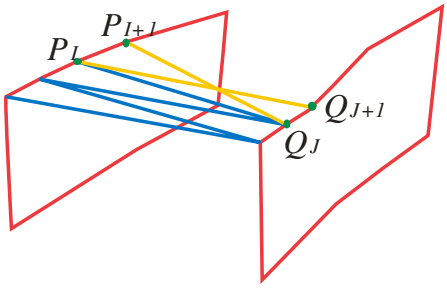


Fig. 1. Connect the adjacent two outlines in space

The detailed instruction of the algorithm is as follows:

(1) For the two stratum curve LP and LQ necessary to be connected, total lengths thereof SP and SQ are respectively calculated, weight counters Φ_P and Φ_Q of the two curves are set to be zero, and current vertex indexes I and J of the two curves are also set to be zero.

(2) If numerical reading of the weight counters Φ_P and Φ_Q is 1, then the connection is finished, quit. Otherwise, execute step (3).

(3) Judge whether the $\Delta P_I P_{I+1} Q_J$ or $\Delta Q_J Q_{J+1} P_J$ is to be connected. Figure 1 indicates the connection selection, and the judgement is made according to value of weight. Calculate length of $\Phi_I = P_I P_{I+1}$ (SP) and length of $\Phi_J = Q_J Q_{J+1}$ (SQ). If $\Phi_P + \Phi_I < \Phi_Q + \Phi_J$, select $\Delta P_I P_{I+1} Q_J$, and set $\Phi_P = \Phi_P + \Phi_I$; otherwise, select $\Delta Q_J Q_{J+1} P_J$, and set $\Phi_Q = \Phi_Q + \Phi_J$. Skip to step (2).

Using the synchronous progress method can effectively connect the adjacent two 2D geological blocks, then the modeling of whole 3D geological blocks can be implemented according to the following method:

(1) Find all 2D geological blocks to be constructed from the existing profiles, and arrange the blocks according to a certain depth direction.
(2) Starting from the first 2D geological block, use the synchronous progress method to connect the adjacent two geological blocks till all 2D geological blocks are connected together, thus the 3D geological block is constructed.

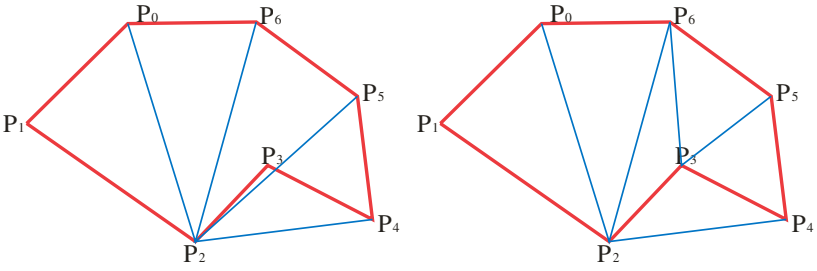
4 The Triangulation and Limited Encryption Algorithm

Delaunay triangulation algorithm is established on the basis of convex hull algorithm, but the surface patch used in the real 3D geological model always requires that the triangulation is concave and even has holes, so it is necessary to limit triangulation; triangulation is implemented according to the limited boundary at the boundary or the internal hole, but not implemented according to standard Delaunay property. Limiting the triangulation includes two steps: at first, limit boundary of the triangulation, such that the triangle in the triangulation complies with the defined limiting boundary strictly so as to avoid intersection of boundary of the triangle with the limiting boundary; at last, delete the redundant triangles from the triangulation according to the defined scope of concave polygon or internal hole.

4.1 Delete the Triangle Outside of the Limited Domain

After embedding limiting boundary, triangles in the triangulation are divided into two sets, triangles in one set completely locate in the limited domain enclosed by the limiting boundary, and triangles in the other set completely locate out of the limited domain. In order to obtain the final limiting triangulation, all triangles out of the limited domain should be deleted. The method of judging whether triangles in the current triangulation are located in the limited domain can be like this: using judgement operator with points in the polygon and substituting one point on the middle line of the triangle to judge.

However, judgement of inclusion relation of a point and a polygon is complex, here is a simpler algorithm. Assume sequence of points in the limiting boundary is anticlockwise, and vertex of each triangle in the constructed triangulation is also arranged anticlockwise. Generally, the two assumptions are easy to meet in the actual use, because sequence of points in a modeling system is always constant and always arranged anticlockwise; for the clockwise sequence, it just needs to reverse result of the algorithm. Because the limited triangles can only be those triangles having limiting boundaries, it is only necessary to judge the triangles having limiting boundaries; sequences of the vertexes of the triangles to be judged are assumed to be V_i, i=1,2,3. Calculate V_0-V_1, V_1-V_2, V_2-V_0, if two of the three formulae have negative results, then the triangle is in the limiting domain; and if only one result is negative, the triangle is not in the limiting domain.



(a) Before (b) After

Fig. 2. Embedding limiting boundary

As shown in Figure 2, the triangle $P_0P_1P_2$ is in the concave domain, because 0-1, 1-2, 2-0, there are two negative results. And the triangle $P_2P_4P_3$ is not in the concave domain, because 2-4, 4-3, 3-2, there is only one negative result. Using this method, only three subtraction operations are necessary to quickly judge whether the triangle is in the limiting domain, so this method is high in efficiency. For the limiting method of the internal holes, the process is just opposite, because purpose of limiting the internal holes is to remove the triangles inside the internal holes, so it just needs to find out the triangles having limiting boundaries of internal holes and carry out the above judgement method reversely.

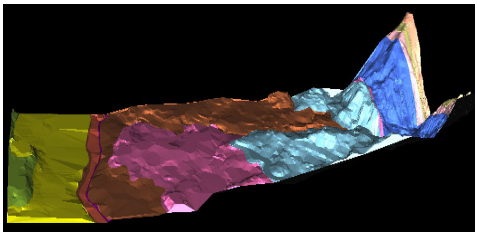


Fig. 3. 3D Surface Model

Figure 3 is the effect drawing of comprehensively using the previous method to carry out triangulation to surface survey data of an work area. We can find that the whole surface is divided into multiple blocks, and different blocks represent different geological attributes, and each is represented by a triangulation; the block situation is complex and many blocks are concave and even have holes. This indicates that the above method of limiting triangulation of triangulation is correct and rational, and is capable of being used in the complex 3D geological modeling.

4.2 Encrypted Triangulation

The triangulation model constructed by using limited triangulation can represent basic geometrical features of surface of geologic body, but the constructed triangulation would be rough and local change thereof is sharp if the control points given by the user are sparse or density of data points is not uniform. In order to construct smoother

3D geological model, it is necessary to encrypt the triangulation by surface interpolation method. The detailed steps of encrypting triangulation are as follows:

(1) Selecting an interpolation method to implement interpolation for the discrete data points, so as to obtain a surface interpolation function;

(2) If there is unprocessed triangle in the triangulation, taking out the triangle and executing step (3);

(3) Checking whether the longest side of the triangle is less than the given threshold value; if surpassing the threshold value, then the triangle needs to be further divided, namely, find the middle point of the longest side and dividing the original triangle into two new triangles, Z coordinate at the middle point is calculated by the previous surface interpolation function; and adding the two new triangles into the triangulation and deleting the original triangle; then returning to the step (2).

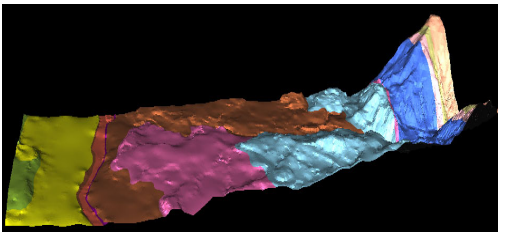


Fig. 4. 3D Surface Model after encryption

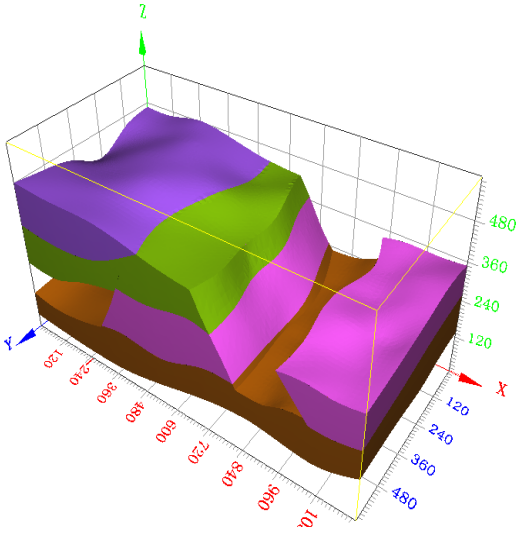


Fig. 5. 3D complex terrain effect drawing

Figure 4 is an effect drawing obtained by implementing interpolation for discrete control points using the B-spline interpolation method and encrypting triangulation by the above algorithm. We can find that the triangulation which is constructed directly

has worse smoothness in the left surface because the control points given by the user are sparse; however, after encryption, the smoothness of the left surface is improved obviously.

5 System Testing and Conclusion

This paper studies deeply the limited triangulation technology used in 3D geological modeling algorithm, and improves the algorithm on the basis of introducing basic incremental insertion Delaunay triangulation algorithm, thus the efficiency of the algorithm is greatly improved such that the triangulation of one million scattered data can be finished within 4-5 seconds. At the same time, in order to realize refined surface description of complex 3D geological blocks, this paper provides a general method of triangulation encryption based on interpolation algorithm. 3D modeling practice shows that the limiting triangulation technology of this paper effectively solves the problems in 3D model representation. As shown in Figure 5. It provides a reliable scientific authority for the decision-making of the underground construction projects etc.

References

1. American National Standards Institute, American National Standard for Information Processing Systems-Computer Graphics-Graphical Kernel System Functional Description, ANSI, X3.124-1985. ANSI, New York (1985)
2. Cerveny, V.: Seismic rays and ray intensities in inhomogeneous anisotropic media. Geophys. J. R. Astron. Soc. 29, 1–13 (2006)
3. Guo, B.: Surface reconstruction from points to spline. Computer Aided Design 29(4), 269–277 (2007)
4. Depreciation: Panel evidence for the G7 and 8 Latin American Countries (April 2005)
5. Hestholm, S.O., Ruud, B.O.: 2-D Finite-difference elastic wave modelling including surface topography. Geophys. Prosp. 42, 371–390 (2004)
6. Angel, E.: Interactive Computer Graphics: A Top-Down Approach with OpenGL, 3rd edn. Edward Angel (2002)
7. Zhao, N., Wen, S.: Based on the profile of complex three dimensional stratum modeling. In: Proceedings of the Third International Symposium on Electronic Commerce and Security Workshops (ISECS 2010), pp. 87–90 (July 2010)

The Teaching Reform and Practice of Business Website Development Course Based on CDIO

Songjie Gong

Zhejiang Business Technology Institute, Ningbo 315012, P.R. China
gsjei2011@163.com

Abstract. At present, the scarcity of engineering talent and quality of engineering education is a common trouble all over the world. Conceive Design Implement Operate (CDIO) is a new engineering education concept and implementation of systems with good prospects for the development and popularization. CDIO teaching idea is applied in course design of business website development by the ubiquitous problem in its traditional practice teaching, and the new course design is carried out and explored. In order to stimulate students interest in learning for starting and training engineering website development ability as a guide, a set of reforms on business website development course are practiced based on CDIO. Through the teaching reform on business website development course of computer specialty for many years, the students' ability of website development has been promoted greatly, and the teaching effect is satisfactory, therefore the goal of modern engineering education reform is achieved.

Keywords: teaching reform; business website development; practice; Conceive Design Implement Operate (CDIO).

1 Introduction

CDIO engineering education mode is an international engineering education reform in recent years. Since 2000, people start research it, after four years of exploration and research, the creation of the CDIO engineering education philosophy, and the establishment of international cooperation to CDIO named organization [1,2,3].

The idea of CDIO is to run the product development life cycle of the product as the carrier, to enable students to active, practical and organic link between courses way to study engineering. CDIO engineering graduates develop the ability to outline the basic knowledge into engineering, personal skills, interpersonal skills and engineering team of four levels of system capabilities, outline requirements for an integrated approach to train students in these four levels of the intended target [4,5]. The CDIO ability is as figure 1 shown.

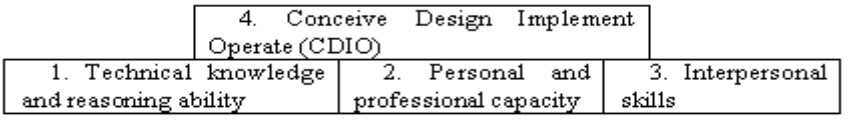


Fig. 1. CDIO Ability

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 508–513, 2011.

CDIO concept of inheritance and development of not only 20 years of engineering education in Europe and the U.S. concept of reform and, more importantly, the system put forward operable capacity-building, and testing the full implementation of the twelve evaluation criteria. So far, dozens of world-renowned universities joined the CDIO organization, and its mechanical systems and aerospace systems, information engineering such as a comprehensive concept of using CDIO engineering education and curricula, and achieved good results, the students trained by deep CDIO model welcomed by the society and enterprises.

2 CDIO Standard

In higher engineering education, the urgent task is to train chinese engineers with international standards, however, the practice of engineering education there are still many problems, such as light weight theory of practice, emphasis on individual academic ability while ignoring the spirit of teamwork, learning the importance of knowledge neglect the cultivation of innovation issues. At present, domestic colleges and universities around the CDIO carried out various forms of seminars, promotion and training activities [6,7,8].

A CDIO program is based on the principle that product, process and system life cycle based development and deployment of the appropriate framework for engineering education. Passionate - design - implementation - operation is a model of the entire product, process and system life cycle. The designing stage includes defining customer needs, taking into account technology, corporate strategy, and regulations, and the development of conceptual, technical and business plans. The design stage focuses on creating the design, that is, the plans, drawings and algorithms that describe what needs to be implemented. The implementation stage relates to the conversion of the design into the product, process or system, including production, programming, testing and validation. The final stage, Operate used the product or procedures implemented to deliver the intended value, including maintenance, developed and retirement of the system [9,10,11]. The standard of CDIO is shown as table 1.

Table 1. CDIO standard

Scale	Criteria
0	There is no plan to adopt the principle that CDIO is the context of engineering education for the program.
1	The need to adopt the principle that CDIO is the context of engineering education is recognized and a process to address it has been initiated.
2	There is an explicit plan to transition to a CDIO context for the engineering program.
3	CDIO is adopted as the context for the engineering program and is implemented in one or more years of the program.
4	There is documented evidence that the CDIO principle is the context of the engineering program and is fully implemented.
5	Evaluation groups recognize that CDIO is the context of the engineering program and use this principle as a guide for continuous improvement.

The product, process and system life cycle is the context for engineering education in that it considers part of the cultural framework or environment in which technical knowledge and other skills taught, practiced and learned. The principle is of a program if there is express consent of the faculty to make the transition to a CDIO program and the support from the program accepted leader sustains reform initiatives.

Starting engineers should be able to understand - Design - Implement - Operate complex value-added engineering products, processes and systems in advanced team-based environments. You should be able to participate in the engineering processes to the development of engineering products, during the work on professional standards in every organization. This is the essence of the engineering profession.

3 Practice of Business Website Development Course Based on CDIO

Systematic teaching program through the design is to achieve implementation of the project approach. Creating innovative practice centers CDIO management of the core issues of teaching, innovative teaching in vocational introduction of international advanced CDIO engineering education philosophy [12,13,14,15]. To "business website development" teaching management reform as the basis, through the systematic design, project system main line, modular courses, and learn the system implementation, flexible management, situational teaching, etc., that will CDIO combined with computer education reform. The full life cycle project management as a carrier, a scientific long-term operation and management of teaching, improve quality of personnel training.

3.1 Systematic Design

In mature based on the integration of teaching, the introduction of international engineering education reform in recent years, the latest achievements of CDIO engineering education model, students in the learning process through concept, design, implementation and operation of product development to product life-cycle operation as the carrier, to enable students to active, practical and organic link between the way curriculum learning business website development.

3.2 Project System Thread

The image of vocational students primarily for the characteristics of thinking to fully learn CDIO the "learning by doing" teaching attitude, focus on allowing students to participate in the project to run the whole life cycle, from project conception, design, production, run four sessions, take the initiative to acquire knowledge, the knowledge associated with the ability to highlight the students hard skills and soft skills training. Therefore, the CDIO curriculum project the main line, set a project, two projects, and three projects. Through the project "Design - making", to consolidate, the use of knowledge and skills, more importantly, focus on the innovative abilities of students; to three items for the comprehensive training, so that students get the hard skills and soft skills further improved.

3.3 The Standard Courses

In the course organization, in order to facilitate the implementation of term project, the implementation of the semester, project-centered curriculum modules. A course module, is a task for the field of field of study curriculum and elective courses with the professional quality, is the project for the term "collage" course, is to use "project is the use of the course", the principle of operation of the project for the semester designed.

3.4 Learn System Implementation

In order to facilitate the implementation of term project, it is to learn organization as a teaching unit. Arrangements for a number of projects, of which the first project idea and design of project-oriented, the second project focused project implementation and operation, the third project focuses on the integrated application. In this way, students is in the completion of the course, through the projects "design - made" so that the integrated using of the knowledge of students, students associated knowledge and capacity capabilities.

3.5 Flexible Management

Implement and learning is a unit of the teaching arrangements, to make the learning cycle time shorter, thus facilitating the production of training or under production of the institute of practice need to be flexible to adjust the contents of each semester, teaching the organization to become flexible.

3.6 The Situation of Teaching

Course is composed of a number of learning situations. Presentation in the teaching situation, the focus on learning process and working process of combining to create a real productive training conditions for students in a typical carrier for each operation, a complete work process in accordance with training programs through systematic repeated training and gradually improve their professional competence. Meanwhile, the teaching situation of the implementation of the six steps using action-oriented approach, the use of project teaching, group work, project reporting, resource development and environmental construction projects and other means to strengthen the soft skills training approach, helping to train high-quality, with sustainable development potential skilled creative talents.

3.7 Combining Learning with the Teaching Model

To train students in independent learning and innovation, respect for individual differences of students to meet the individual needs of the development. Of the more outstanding students, the students learn there is spare capacity, abilities, can require the introduction of enterprise application project, under the guidance of students in teachers studio project development, project and by the school teachers and business assessment team composed of experts in assessment evaluation, evaluation results on the performance as a curriculum to train students in practical work. Students are in

studio and the mechanism is essentially an entity, and it has two important functions: first, students develop mechanisms; the other is for the community, business unit service mechanism. In students learning, there are spare capacity-building activities to participate in the talented students, the general direction of two training: one is preparing to practice all kinds of software design competition, and the other is a combination of, involved in the actual school-enterprise cooperation, services, and social services directly.

4 Conclusions

At present, the shortage of engineering talent and quality of engineering education is a common trouble all over the world. CDIO is a new engineering education concept and implementation of systems with good prospects for the development and popularization. CDIO teaching idea is applied in course design of business website development by the ubiquitous problem in its traditional practice teaching, and the new course design is carried out and explored. In order to stimulate students interest in learning for starting and training engineering website development ability as a guide, a set of reforms on business website development course are practiced based on CDIO. Through the teaching reform on business website development course of computer specialty for many years, the students' ability of website development has been promoted greatly, and the teaching effect is satisfactory, therefore the goal of modern engineering education reform is achieved.

References

1. Benchmark table for first-year courses. CDIO Project document. Chalmers University of Technology (2002)
2. Conover, W.J.: Practical Nonparametric Statistics, 2nd edn., pp. 130–133. Wiley & Sons, Chichester (1980)
3. Crawley, E.: The CDIO Syllabus: A Statement of Goals for Undergraduate Engineering Education. MIT CDIO Report #1, Dept. of Aeronautics and Astronautics. Massachusetts Institute of Technology, Cambridge, MA, USA (2001)
4. Gustafsson, G., Malmqvist, J., Newman, D.J., Stafström, S., Wallin, H.P.: Towards a new model for first-year introductory courses in engineering education programmes. In: NordDesign 2002, Trondheim, Norway, August 14-16 (2002)
5. Hurst, K., James, R.D., Raines, M.: The Progressive Use of Projects in an Engineering Degree Course. The International Journal of Mechanical Engineering Education 21(4), 373–379 (1993)
6. Newman, D.J., Amir, A.R.: Innovative First Year Aerospace Design Course at MIT. ASEE Journal of Engineering Education 90(3), 375–381 (2001)
7. Newman, D.J.: Interactive Aerospace Engineering and Design. In: Introductory Textbook with Accompanying CD-ROM Multimedia. McGraw-Hill, New York (2002)
8. West, H.: A Criticism of an Undergraduate Design Curriculum. Design Theory & Methodology, DE- 31, 7–12 (1991)
9. Shell Energy Scenarios to 2050 (2050), http://www.shell.com/home/content/aboutshell/our_strategy/shell_global_scenarios/shell_energy_scenarios_2050/shell_energy_scenarios_02042008.html

10. Lettinga, G., Zeeman, G., Kujawa, K., Mels, A., van Lier, J.: Sustainability in Environmental Protection for all citizens, http://www.ecosanlac.org/ecosanlac/assets/lettinga.pdf
11. Crawley, E., et al.: Rethinking Engineering Education. Springer, New York (2007)
12. Edstrom, D., et al.: Teaching and Learning. In: Crawley, E.F. (ed.) Rethinking Engineering Education: The CDIO Approach. Springer, New York (2007)
13. Kelly, H.G., Urban, W., Warren, R.: Design Considerations for Autothermal Thermophillic Aerobic Digestion (2003), http://cedb.asce.org/cgi/WWWdisplay.cgi?0521907
14. 2PAD Digester System, http://www.degremont-technologies.com/dgtech.php?article389&gclid=CMDNj-yv9ZcCFcMupAodSHnIDA
15. Wells, G., Claxton, G.: Learning for Life in the 21st Century. Blackwell Publishing, Oxford (2002), ISBN 0-631-22330-4

A User-Based Collaborative Filtering Recommendation Algorithm Based on Folksonomy Smoothing

Feng Ge

Ningbo City College of Vocational Technology, Ningbo 315100, P.R. China
gefengei@sina.com

Abstract. Recommender systems suggest useful and interesting products to customers in order to increase customer satisfaction and online conversion rates. Collaborative filtering recommendation algorithm is the most usually applied recommender system for personalized. In view of the fact that collaborative filtering systems depend on neighbors as information sources, the recommendation quality of collaborative filtering relies on the neighbors selected. However, traditional collaborative filtering has some essential limitations in selecting neighbors. One of these chiefly problems is data sparsity. While the number of items is increase, the ratio of common rated items is decrease so calculating the computations of neighborhood become difficult. To alleviate the sparsity, a user-based collaborative filtering recommendation algorithm based on folksonomy smoothing is presented. The approach firstly fills the empty using folksonomy technology. And then produce the recommendations employing the user-based collaborative filtering algorithm.

Keywords: collaborative filtering; recommendation algorithm; folksonomy; smoothing.

1 Introduction

In the increasingly competitive environment, in order to effectively preserve the customers and increase sales of e-commerce systems, e-commerce recommendation system has been revealed. Product recommendation systems in e-commerce system provide commodity information and advice to help customers decide what products to buy and analog sales staff recommends products to customers complete the purchase process, so that customers feel completely personalized service [1,2].

Currently, the most personalized recommendation technology that used in filtering algorithms is the user-based collaborative filtering method [3]. The user-based collaborative filtering recommendation algorithm is for customers to purchase products and the current system data in the other purchased this product with other customers purchase and evaluation of related products. Collaborative recommended starting point is, the user may prefer to buy their products that have similar products in the recommendation process, user can evaluate a variety of merchandise to determine the degree of similarity between products, and then recommend those closest interest products to the user.

With the development of electronic commerce, e-commerce sites increasing the number of products, the user-based collaborate filtering gradually exposed its

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 514–518, 2011.

shortcomings: firstly, the data sparsity problem, that is a huge site number and customer purchases of products or to make the evaluation only a small part of the opposition. The sparse data accuracy will lead to the reduction recommendation algorithm.

2 Smoothing Based on Folksonomy

2.1 Folksonomy

Folksonomy [4,5,6,7,8] is a free choice of the keywords the user to collaborate on the site classification way. Folksonomy can also be seen as one from the crowd that can easily see that the present concern the subject matter and the corresponding way. Taxsonomy is a simple form of classification. Listed on the site can be seen as a simple classification according to the form of various types. Compared with the taxsonomy, folksonomy is a classification of liberalization. It is not a direct classification by the manufacturer site, but by the general public category. All users can order their own subjective classification of related content. These users can know their labels also be related to the content.

2.2 Tags Matrix as TFxIDF

The set of users U = {u1, u2, …, ui, …, um}, the set of items I = {i1, i2, …, ij, …, in}, the set of tags D = {d1, d2 , … dk, …, dl}. Score matrix M = (Rij) m × n, where Rij is the rating of user ui to item ij.

TFxIDF is a weight, commonly used in information retrieval and text mining, using a statistical method to assess a collection of documents in a word on the importance of a document. Here, the weights are used for similarity calculation.

If the user uses the label as a document, get a weight TFxIUF, so that users get the weight of the preference of a label and the matrix similar to the rating matrix. It can calculate the similarity between users.

2.3 Calculating TFxIDF

The user the set of ui labels SDui, the user ui tag number of all projects recorded as CDui (A), the user ui dk marked with labels denote the number of projects CDui (k), then the label dk all the labels in the user ui. The frequency of Fui (k) can be calculated as following.

$$Fui(k) = \frac{CDui(k)}{CDui(A)}$$

Definition NF (k) is the inverse user frequency tags of dk, with the following formula.

$$NF_k = \log \frac{CD(A)}{CD(k)}$$

Which, CU (A) = | U | means that all users, CU (k) = | {ui | dk ∈ D} | dk said the number of users have tagged.

Defined in the user ui label dk in TFxIUF frequency tf × iuf (k, i), and tf × iuf (k, i) = Fui (k) • NFk, a matrix MM1, expressed as the following formula.

$$MM1 = \begin{bmatrix} s(1,1) & s(1,2) & \cdots & s(1,q) \\ s(2,1) & s(2,2) & \cdots & s(2,q) \\ \vdots & \vdots & \ddots & \vdots \\ s(p,1) & s(p,2) & \cdots & s(p,q) \end{bmatrix}$$

Where, s (i, j) = tf × iuf (i, j).

2.4 Tags Similarity

In this section, we use the Pearson correlation coefficient to calculate the similarity between users based on the tag weight matrix. It is different from the score matrix based on scores calculated by the similarity weight matrix-based label similarity as calculated by the label similarity.

The tag similarity of users is using the following formula.

$$SIM(r,s) = \frac{\sum_{i \in Srs} (tu(i,r) - A(r))(tu(i,s) - A(s))}{\sqrt{\sum_{i \in Srs} (tu(i,r) - A(r))^2} \sqrt{\sum_{i \in Srs} (tu(i,s) - A(s))^2}}$$

2.5 Smoothing

The prediction score of target user Ui on the target item Ij is calculating by the following formula.

$$Pij = \frac{\sum_{y \in In} RiySIM(j,y)}{\sum_{y \in In} SIM(j,y)}$$

3 Prediction Employing User-Based Collaborative Filtering

The collaborative filtering recommendation method based on users can be divided into three steps: data representation, forming the neighbors of target user and producing the recommendations [9,10,11]. Data representation is a modeling problem that the users buy the products. Forming neighbor's task is to study the formation of a neighbor how to identify the users. Producing recommended for users is to find the top N products from the neighbors selected.

3.1 Data Representation

In a typical collaborative filtering-based recommendation systems, user input data is a collection of transactions of products, and usually can be represented as a m × n user - product matrix as R.

Each user is represented by item-rating pairs, and can be summarized in a user-item table, which contains the ratings Rij that have been provided by the ith user for the jth item, the table as following.

Table 1. User-item ratings table

Item / User	Item1	Item2	… …	Itemn
User1	R11	R12	… …	R1n
User2	R21	R22	… …	R2n
… …	… …	… …	… …	… …
Userm	Rm1	Rm2	… …	Rmn

Where Rij denotes the rating of item Ij rated by a user Ui. If user Ui has not rated item Ij, then Rij =0. The symbol m denotes the total number of users, and n denotes the total number of items.

3.2 Forming Neighbors

Collaborative filtering-based recommendation system, the most important step is to target users based on similarity calculation of neighbor users. The main objective of the formation of a neighbor, for a user u, to find an orderly queue l users N = {N1, N2, ···, Nl}, makes the sim (u, N1) of the value of the maximum, sim (u, N2) The value of the second, and so on. Similarity measure commonly used functions are Pearson correlation coefficient and Cosine measure.

Pearson's correlation measures the linear correlation between two vectors of ratings, as shown in following.

$$SIM2(i,j)=\frac{\sum_{c\in I_{ij}}(R_{ic}-A_i)(R_{jc}-A_j)}{\sqrt{\sum_{c\in I_{ij}}(R_{ic}-A_i)^2\sum_{c\in I_{ij}}(R_{jc}-A_j)^2}}$$

3.3 Prediction

Once the target user's neighborhood, you can recommend the following two forms:

1. Predict a user likes a product level. Generally calculated using the following formula:

$$P_{ut}=A_u+\frac{\sum_{i=1}^{c}(R_{it}-A_i)SIM2(u,i)}{\sum_{i=1}^{c}SIM2(u,i)}$$

2. Referring users to a range of products, which is well-known top-N recommendation.

4 Conclusions

Recommender systems suggest useful and interesting products to customers in order to increase customer satisfaction and online conversion rates. Collaborative filtering systems depend on neighbors and the recommendation quality of collaborative filtering relies on the neighbors selected. However, traditional collaborative filtering has some essential limitations in selecting neighbors. One of these chiefly problems is data sparsity. While the number of items is increase, the ratio of common rated items is decrease so calculating the computations of neighborhood become difficult.

To alleviate the sparsity, in this paper, we presented a user-based collaborative filtering recommendation algorithm based on folksonomy smoothing. The approach firstly fills the empty using folksonomy technology. And then produce the recommendations employing the user-based collaborative filtering algorithm. At last, we can get the top N recommendation.

References

1. Gao, F., Xing, C., Du, X., Wang, S.: Personalized Service System Based on Hybrid Filtering for Digital Library. Tsinghua Science and Technology 12(1), 1–8 (2007)
2. Huang, Q.-H., Ouyang, W.-M.: Fuzzy collaborative filtering with multiple agents. Journal of Shanghai University (English Edition) 11(3), 290–295 (2007)
3. Lekakos, G., Giaglis, G.M.: Improving the prediction accuracy of recommendation algorithms: Approaches anchored on human factors. Interacting with Computers 18, 410–431 (2006)
4. Dotsika, F.: Uniting formal and informal descriptive power: Reconciling ontologies with folksonomies. International Journal of Information Management (29), 407–415 (2009)
5. Kome, S.H.: Hierarchical subject relationships in folksonomies. Master's thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, vol. (11) (2005)
6. De Meo, P., Quattrone, G., Ursino, D.: Exploitation of semantic relationships and hierarchical data structures to support a user in his annotation and browsing activities in folksonomies. Information System 34, 511–535 (2009)
7. Chen, W.-H., Cai, Y., Leung, H.-F., Li, Q.: Generating ontologies with basic level concepts from folksonomies. In: International Conference on Computational Science, ICCS 2010 (2010)
8. Gruber, T.: Ontology of Folksonomy: A Mash-up of Apples and Oranges. AIS SIGSEMIS Bulletin (2), 3–4 (2005)
9. Degemmis, M., Lops, P., Semeraro, G.: A content-collaborative recommender that exploits WordNet-based user profiles for neighborhood formation. User Model. User-Adap. Inter. 17, 217–255 (2007)
10. Yu, L., Liu, L., Li, X.: A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce. Expert Systems with Applications 28, 67–77 (2005)
11. Herlocker, J.: Understanding and Improving Automated Collaborative Filtering Systems. Ph.D. Thesis, Computer Science Dept., University of Minnesota (2000)

Study about the Model of Business Production Logistic and Optimizing*

Guojing Xiong and Minjian Zhou

School of Economy and Management Nanchang University Nanchang, Jiangxi Province, China
Graduate School Nanchang University Nanchang, Jiangxi Province, China
xiongguojing@ncu.edu.cn, zmj2696@sina.com

Abstract. The design, plan and management of good production logistics are the keys to improve labor production quota and produce quality. Only reasonable production logistics of organization can make the process of business production always in the best state. This article that analyzing the main factors and shortcoming that influence business production. Discussed the way that solve the problems and obstacles of business production logistics, constructed the model of business production logistics system based on lean logistics, put forward the tactics that optimize production logistics system, reduce the logistics cost, renew the management sense of business production logistics, improve the key advantage of business.

Keywords: Production Logistics, Model, Lean Production, Supply Chain Management.

1 Introduction

Production logistics is the logistics activities in internal production process which responsible for transportation, storage, loading and unloading of materials and other tasks, it's the foundation of the continued product and constitutes the whole process of the internal logistics activities. The management of Production logistics should meet the requirements in order to ensure stable production and coordinated manner, shortening the production cycle, improves product quality and reduces product consumption.

Production logistics is a part of the production method, there are many relationship between them, some of the logistics process to implement processing and manufacturing which required by the production technology; some of them are when in the process of manufacturing it complete the logistics at the same time; some others are link with different manufacturing segments by the logistics. Logistics occupant and consume of resources is an important part of the cost of production. Logistics activities are frequently that we have in production, so the production logistics have great impacted on production costs.

* This work is supported by the Jiangxi Province Office of Education Social Science Foundation Grant # JC0918.

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 519–526, 2011.

Production logistics should meet the requirements of specialized production to face the specific logistics needs instead of facing the society and the general logistics needs. Therefore, logistics achieve higher efficiency by the means of specialization and standardization, and would adjust the scale of business to the changing clients. Thus we can implement our planning more accurate and precise and using the effective means of resource management systems to achieving lean logistics.

2 The Main Factors and Shortage of Influence the Production Logistics

2.1 Production Logistics Management Is an Important Segment of the Supply Chain Management

It's based on the optimized resources and capabilities to produce the best product in a lowest cost, fastest speed and quickly to meet user's requirements such as variety, quality, quantity and delivery in order to improve enterprise responsiveness and efficiency as well as reduce non-value-added services. It started when companies accept orders including contract processing, organize to apply raw materials, the production planning, manufacturing, formulation and issued the command, control and adjust of the production process, production and performance of collection and sorting until the process of tissue products manufactured.

2.2 The Main Factors of Production Logistics

Type of production: the product variety, structure, complexity, accuracy class, process requirements and raw materials are different from production types. These characteristics affect the composition of production logistics and the ratio between each other.

The scale of production: The scale of production is the output of products per unit time(usually expressed in annual output).While the Production scale is larger, more complete composition of the production process, the greater the amount for the logistics. On the contrary, while production scale is small, there are no very fine conditions to divide composition of the production process small as well as small of the logistics capacity.

The level of business professionalism and collaboration: as the raising level of social specialization and collaboration, the enterprise will tend to streamline the production process and shortened the logistics process. Thus some of the basic stages of the process semi-finished products such as rough, spare parts, components, etc., can be supplied by the other specialized factories.

2.3 The Main Shortage of the Production Logistics

(1) Infrastructure, technology and equipment level falls behind. (2) The sharing of information of network is in a low level, information transmission lag. (3) The market mechanism is not perfect and the order of competition is not standardized, the existing rules and regulations and management is not adapted to the modern logistics development. (4) The concept of supply chain management is weak; the degree of logistics standardization is not high. (5) Inventory management level is low; except

for a few enterprises realize zero inventory management, the production and logistics management are still not high. (6) The inventory takes up a lot of money, production and logistics cost accounting unreasonable.

3 The Problems, Obstacles and Solutions of the Production Logistics

3.1 The Main Shortage of the Production Logistics

(1) Infrastructure, technology and equipment are backward, lack the total and low utilization are co-exist, facilities and equipment are lack of effective diversity and have functional defects. (2)The inadequate of production logistics are low level of information and lack of professionals (3) China's logistics market is still in its infancy, the market mechanism is not perfect, the order of competition is not standard and the existing rules and regulations and management is not adapted to the modern logistics development. (4) Theoretical research and the actual operation of modern logistics are lagging behind (5) Inventory management is not high. Except for a few enterprises in achieving zero inventory management, most manufacturers are still not high inventory levels. Inventory take up a lot of money and goods are aging, deterioration and other issues.

3.2 The Main Shortage of the Production Logistics

(1) The concept of Supply Chain Management (SCM) is weak, supply chain nodes among enterprises are lack of cooperation and coordination. (2) The level of network sharing of information is low, and information transfer is delay (3) Logistics is not high degree of standardization (4) Production logistics management is decentralization, and the level of social services is low. Because of the separation of departments, lack of unified planning and redundant construction intensified, National Logistics Company are in the small, many, scattered and weak condition, so it is difficult to form effective social service network. (5) Production logistics costing is unreasonable.

3.3 The Solutions of Production Logistics

Strengthen the logistics infrastructure construction and reconstruction: the modernization of logistics infrastructure is the hardware to achieve logistics modernization, so enterprises ,under the conditions of funding, should strengthen the logistics infrastructure investment, establish a number of automated warehouse, update transportation equipment and purchase of computer hardware and software for logistics information processing and so on.

Improve inventory management level and realize enterprise inventory management: we can be used to order production, that is to say, all production activities of enterprises are carried out according to the order procurement, manufacturing, distribution, and warehouse logistics operation is a site. Or by way of JIT delivery, that is to say enterprise will delivery produced goods in accordance with the order to customers hand timely. So enterprises can reduce inventory by goods in transit transportation and distribution processing.

Improve the index system optimization of production logistics and production planning optimization system: regulate production and planning process, market forecast varieties and yield optimization can be to solve the problems of production planning and production scheduling.

Achieve systematic management of production logistics: production logistics emphasis on modern logistics system as a whole. From the concept of systematic go to design, analysis, research, and improve the production logistics system and pursuit the overall system optimization and efficient.

Make logistics costing clarification: enterprises are not only having a unified logistics management but also a clear logistics cost to go to quantify the analysis and evaluate logistics efficiency, only these enterprises can achieve production logistics enterprise objectives.

Strengthen the capabilities of logistics information: information processing is an effective guarantee of smooth logistics activities. Information processing technology is the core of information processing capability, information processing technology is the premise of supply chain establishing a rapid response strategy. Improving the logistics information processing capacity can better enhance the core competitiveness.

4 The System Model of Production Logistics and Optimization

4.1 Enterprise Production Logistics Based on Lean Logistics

Lean Logistics uses lean idea into the management of logistics activities. The core of lean thinking is the elimination of all waste, which is providing the right amount of product at the right time, right place. It stresses an idea which is the customer first, timely and accurate, the overall optimization, continuous improvement and innovation. Lean production Logistics is the concrete manifestation of lean production in the process of enterprise manufacture activity, is based on the theory of lean manufacturing requirements of the production logistics, production logistics, no waiting, no continuity of current flow, thereby reducing the waste of the production process and shorten the production cycle, reduce production costs and improve efficiency, it is the basis for achieving lean logistics. Lean Logistics on the establishment of the production logistics system helps to extensive changes in the management of manufacturing concepts helps to improve core competitiveness of enterprises. Information logistics system is the key to achieving lean logistics and is a prerequisite for the implementation of lean logistics.

4.2 Model Based on the Lean Logistics of Enterprise Production Logistics

In lean logistics system, the demanding of customers is the driving force of production source, is the starting point of value. Lean production logistics system provides stricter requirements to Production Logistics system. It Requires production logistics has certain flexibility and adaptability to the environment, agility, and punctual delivery, high reliability, fast response to the market, reducing inventory costs, etc. Therefore, enterprise will change traditional production logistics operation model into lean production logistics operation model, from the user requirement and requirement of information flow to reverse pull production logistics flow. In the

production process, the next working procedure as customers, by the next working procedure demand instructions to perform production and supply. Follow the principle, according to the needs of customers JIT the time, place, quantity, quality and price production and supply products, the realization of zero inventory, zero defect and zero fault production logistics services. (Figure 1)

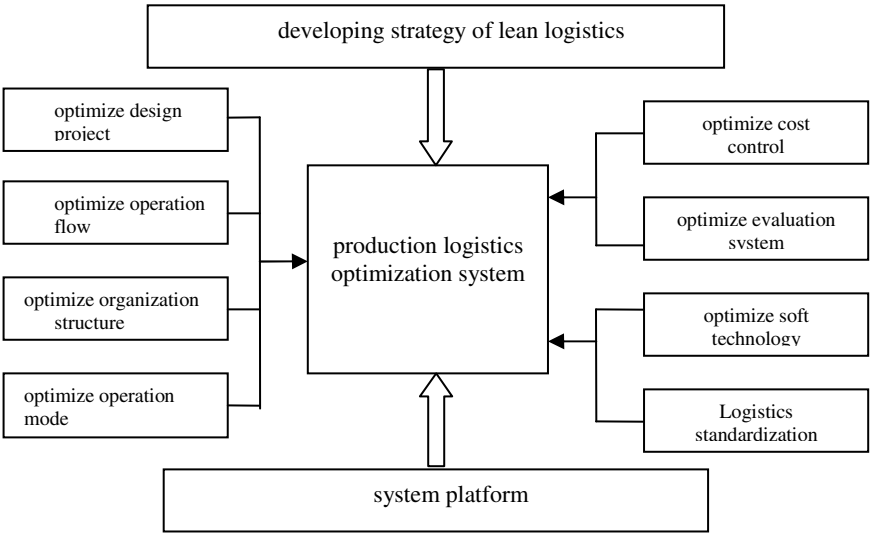


Fig. 1. Model of production logistics system based on the lean logistics

4.3 Enterprise Manufacturing Logistics System Optimization Strategy

Enterprise production logistics operation mode optimization strategy. Production and logistics control is based on the production scheduling , dynamic control and adjustment in the whole production and logistics process from raw materials to the plant products factory, ensure to finish production plan and the provisions of the contract product punctually, quality, quantity. Strengthening production and logistics process control solve the production scheduling optimization preparation, logistics tracking, production scheduling, etc.

Enterprise production logistics cost management optimization strategy. Enterprise production logistics cost management goal is minimum inventory, integrating transportation, assure product quality, etc. Implement manufacturing logistics cost accounting management; make the production logistics cost accounting specialization, the exposition. To accurately calculate production logistics costs, the first to make base data finishing work. Second, must be clear about the composition of the logistics cost. Then, to production logistics cost analysis and comparison with manufacturing costs and other expenses, the relationship between be mastered logistics system costs.

Enterprise production logistics standardization and management technology optimization strategy. Logistics standardization is promoting lean logistics base, is China's logistics and international strong foundation. Enterprises as logistics standards implementation unit should cooperate with standards and conscientiously implement

the national standards. Implementation of logistics standardization management, not only the dimension and logistics of logistics operation procedures and methods should as far as possible standardization and logistics performance evaluation system also should be standardized.

Enterprise production logistics performance appraisal system optimization strategy. Logistics performance evaluation for reducing logistics cost, improve the management efficiency has a pivotal role. Accordingly, we should focus on the strengthening of the enterprise culture construction and to improve the staff of logistics performance management knowledge, and logistics performance evaluation result with employee's reward hooks. To strengthen logistics budget management, cost management and risk management.

4.4 The Strategy of Enterprise Manufacturing Logistics System Optimization

To implement material handling distance shortest configuration. Within the workshop of decorate and the reasonable decorate each workshop of logistics lines plays a decisive role. Therefore, reasonable decorate each workshop and within the workshop of each process position, avoid material handling in various processes between circuitous handling, round-trip handling, backflow handling. And try to improve handling the batch and reduce the number of handling, make handling distance retreat to a minimum.

The use of reasonable production mode. In the early period meet production and users of prophase and tried to compress inventory, reduce the intermediate in-process inventory of storage, had better be completely just-in-time production. Namely, the former procedure warmly machined out of semi-finished product quantity just is the next procedure demand, coordinated and cancel middle inventory, reduce transport handling, according to the number of materials and the shapes of the nature, material flow rate of the size and distance to apply the proper handling equipment and handling methods.

Enterprise workshop reasonable decorates. First, the enterprise internal workshop, position and within the workshop working procedure, and the workshop layout of the finished product area, semi-finished area, excipient area as well as other relevant auxiliary facilities position distribution. Reasonable decorate can make the material or product handling the minimum amount. Secondly, logistics line is determined thing, big part of the problem is to choose the appropriate handling equipment. The equipment should be able to adapt to be handling of material nature, weight, and shape, size and material flow rate. Both must make equipment fixed investment is less and make its operation cost is small.

Coordination and balance production logistics bottleneck link. Adopt optimized production logistics management method, to bottleneck resources plan is analyzed and more effective coordinated project plan, which is beneficial to the key project plans, at the same time, also can overcome pure traditional production logistics in a bottleneck resources a possible ability waste. In addition, through to the bottleneck resource capacity balance, discovered early capacity, in order to take measures, gaps assure integral production plans.

Improve production logistics control efficiency. In the integrated system, mobilize workers' enthusiasm and creativity, the production process control and harmonious by

mutual is directly linked to finish production workers. The dispatcher only when necessary, such as how to change the process route, change processing equipment, that is executed control and coordination function, this helps to reduce the complexity of dispatching work, improve the scheduling the quality of work, so as to raise productivity.

5 Conclusions

Production logistics is the core business, only it has the rational organization of production logistics, it can make the production process always at its best. Production logistics system is an important part of the production process and it is the link of contacting all sectors of manufacturing into an organic whole. Good production logistics design, planning and management are the key to improve productivity and product quality. Logistics information system plays the nerve center role of logistics activities, it can improve logistics management system, achieve production logistics functions the transition from functionalization to integration and establish enterprise logistics management information system, and the system can fully observe and control the entire logistics system operation, and it also can achieve dynamic management. Reasonable allocation of manpower, material resources can make allocation of resources more optimal and make the relevant parts no longer waiting for each other, so this can avoid flow stagnation.

Production logistics management should begin from standardizing enterprise management data, it carries out real-time monitoring and precautionary during production logistics process. So the problems can be found in the shortest time to correct and remedy. Based on the basic data, establishing a reasonable production management plan, optimizing production scheduling, strengthening production logistics process control can make logistics, cash flow and information flow of the logistic activities in the best state, thus achieve significant economic benefits and social benefit.

Establishing a customer-centric, process-oriented organizational structure, and applying lean logistics which highlights customers first, timely and accurate, the overall optimization, continuous improvement and optimization of production and logistics systems go to achieve optimized production logistics organization. Production logistics management and control is conducive to optimize production logistics, reduce production costs and improve product competitiveness. With suitable suppliers, storage and transportation business forming strategic alliances and taking participation in the domestic and international competition by the overall advantages of the supply chain this can improve product competitive ability and market share.

References

1. Duan, C.-J., Liu, M.-X.: Selection of Logistics Models Based On Competitive Strategy. Logistics Engineering and Management 31(180), 26–32 (2009)
2. Li, Y.-F., Liu, W.-S.: Logistics Planning Based on Industry Supply Chain. Commercial Research (400), 207–209 (2010)

3. Li, Y.-S., Jiang, W.-S.: Logistics System and its Reconstruction Model Based on Agile SC. Logistics Technology 27, 41–43 (2008)
4. Shen, W., Ma, S.-H.: A Scheduling Model for Logistics Capability in A Time-based Multi-stage Supply Chain. Operations Research and Management Science 16(3), 20–25 (2007)
5. Zheng, J.-C., Li, Y.-S.: Tactics Selection on Lean M.F and Agile M.F. Journal of Beijing Jiaotong University 3(12), 29–32 (2004)
6. Zheng, P.-F., Tian, Q.-H.: Research on the Typical Production Logistics System and Modeling Technology vol. 20(5), pp.148–152 (2006)
7. Zheng, Y.-Q., Zhang, X.-Y.: The Lean Improvement Method and Application Based on Production Process Analysis. Manufacture Information Engineering of China 38, 18–22 (2009)

Analysis of Single-Phase APF Overtone and Idle Current Examination

Li Yang[1] and Xin Ning[2]

[1] Department of Electrical and Automatic Engineering,
Nanchang University Nanchang, China
[2] Department of Electrical Engineering,
Liming Vocational University Quanzhou, China
{yangliedu,nxdoudou}@163.com

Abstract. This paper introduced three kind of single-phase APF overtones and idle current examination method first,then compare with MATLAB simulation based on these methods,receives a satisfactory conclusion.

Keywords: single-phase; APF; idle current.

1 Introduction

In recent years, along with the misalignment electric power electronic device is widely applied day by day, the electrical network has produced the massive overtones. The electricity iron load is the typical misalignment overtone source, the Zhejiang and Jiangxi electric railway's clear the electrical network has brought certain influence for Jiangxi, therefore, it is necessary to carry on the analysis to the Zhejiang and Jiangxi iron overtone, and proposed that related suppressed the strategy. This paper studies is suitable for electric railway's single-phase APF[1], carries on the analysis to its harmonic current examination method.

2 Instant Reactive Power Single-Phase Overtone Idle Work Examination Method

The inspection based on the instant reactive power theory's three-phase circuit harmonic current examination method, discovered will always examine the first three-phase signal becomes the mutually perpendicular $\alpha\beta$, then further calculates again. Regarding the single-phase circuit, simplifies above method, only need construct interphase current and the actual electric current lags T/4 (here T is again power frequency cycle), the direct production supposition's two phase coordinates are signals. [2]

Supposes the network voltage spurt value is:

$$u_s(t) = \sqrt{2}U \sin \omega t \tag{1}$$

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 527–534, 2011.

The electrical network current spurt value is:

$$i_s(t) = \sum_{n=1}^{\infty} \sqrt{2} I_{2n-1} \sin[(2n-1)\omega t + \varphi_{2n-1}] \tag{2}$$

In $\alpha\beta$ coordinate system, designated $i_s(t)$ for α signal, lag T/4 it to construct β signals, obtains the imaginary $\alpha\beta$ signals, as follows:

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^{\infty} \sqrt{2} I_{2n-1} \sin[(2n-1)\omega t + \varphi_{2n-1}] \\ \sum_{n=1}^{\infty} \sqrt{2} I_{2n-1} \sin[(2n-1)\omega (t - T/4) + \varphi_{2n-1}] \end{bmatrix} \tag{3}$$

According to the reactive power theory, may calculate:

$$\begin{bmatrix} i_p \\ i_q \end{bmatrix} = \begin{bmatrix} \sin\omega t & -\cos\omega t \\ -\cos\omega t & -\sin\omega t \end{bmatrix} \begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \begin{bmatrix} \bar{i}_p \\ \bar{i}_q \end{bmatrix} + \begin{bmatrix} \tilde{i}_p \\ \tilde{i}_q \end{bmatrix} = \begin{bmatrix} \sqrt{2} I_1 \cos\varphi_1 \\ \sqrt{2} I_1 \sin\varphi_1 \end{bmatrix} + \begin{bmatrix} \tilde{i}_p \\ \tilde{i}_q \end{bmatrix} \tag{4}$$

Above various in the formula, $\sin\omega t$ and $\cos\omega t$ are the voltage signal which obtains after the zero crossing synchronization and the phase-locked loop. Above the variable "-" the expression direct component, "~" the expression AC component, $\bar{i}_p$ and $\bar{i}_q$ express the electric current separately active and the idle work direct component. The inverse transformation may obtain the fundamental current $i_{\alpha f}$ and $i_{\beta f}$ [3]。 .

$$\begin{bmatrix} i_{\alpha f} \\ i_{\beta f} \end{bmatrix} = \begin{bmatrix} \sin\omega t & -\cos\omega t \\ -\cos\omega t & -\sin\omega t \end{bmatrix}^{-1} \begin{bmatrix} \bar{i}_p \\ \bar{i}_q \end{bmatrix} = \begin{bmatrix} \sqrt{2} I_1 \sin(\omega t + \varphi_1) \\ \sqrt{2} I_1 \sin(\omega t - T/4 + \varphi_1) \end{bmatrix} \tag{5}$$

Subtracts the fundamental wave component from the full current,then obtain the harmonic current component:

$$i_{sh} = i_\alpha - i_{\alpha f} \tag{6}$$

Based on the instant reactive power's single-phase overtone examination control diagram, as shown in Figure 2.1. LPF is the low pass filter, PLL is the network voltage synchronization phase-locked loop[4], the cosine signal has the electric circuit combined action to produce and the network voltage which in transformation matrix C needs with the phase sinusoidal signal and the corresponding cosine signal. When separates the q channel, after examining the overtone idle work adds together, then simultaneously examines the harmonic current and the idle current.

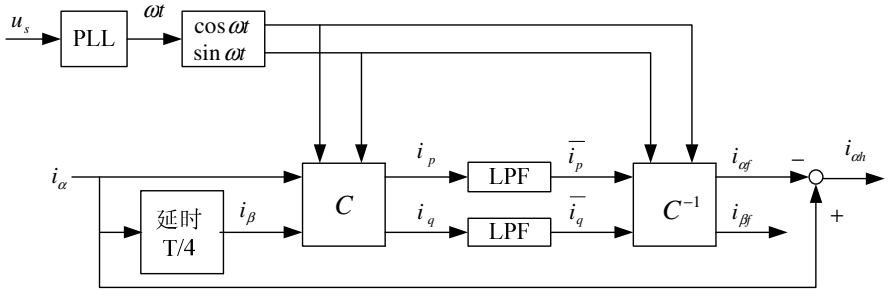


Fig 2.1. Based on instant reactive power single-phase overtone examination control diagram

3 Wattful Current Separation Overtone and Idle Current Examination Law

In this method, uses with the network voltage frequency unit cosine, the sinusoidal signal multiples directly separately with the electrical network electric current, and after low pass filter, obtains in the electrical network electric current instantaneous fundamental wave wattful current and the instantaneous fundamental wave idle current, then obtains the instant harmonic current.

Supposes the network voltage same type (1), the electrical network electric current is:

$i_q = \sqrt{2} I_q \sin\varphi\cos\omega t$ is electrical network instant fundamental wave idle current

$i_h = \sum_{n=2}^{\infty} \sqrt{2} I_{2n-1} \sin[(2n-1)\omega t + \varphi_{2n-1}]$ is electrical network instant harmonic current

$$i_s(t) = \sqrt{2} I_1 \sin(\omega t + \varphi) + \sum_{n=2}^{\infty} \sqrt{2} I_{2n-1} \sin[(2n-1)\omega t + \varphi_{2n-1}] =$$

$$\sqrt{2} I_1 \cos\varphi \sin\omega t + \sqrt{2} I_1 \sin\varphi \cos\omega t + \sum_{n=2}^{\infty} \sqrt{2} I_{2n-1} \sin[(2n-1)\omega t + \varphi_{2n-1}] = \quad (7)$$

$$\sqrt{2} I_p \sin\omega t + \sqrt{2} I_q \cos\omega t + \sum_{n=2}^{\infty} \sqrt{2} I_{2n-1} \sin[(2n-1)\omega t + \varphi_{2n-1}] = i_p + i_q + i_h$$

The type (7) nearby two simultaneously is multiplied by $2\sin\omega t$:

$$2 i_s(t) \sin\omega t = 2\sqrt{2} I_p \sin^2 \omega t + 2\sqrt{2} I_q \cos\omega t \sin\omega t +$$

$$\sum_{n=2}^{\infty} 2\sqrt{2} I_{2n-1} \sin[(2n-1)\omega t + \varphi_{2n-1}] = \sqrt{2} I_p - \sqrt{2} I_p \cos 2\omega t + \quad (8)$$

$$\sqrt{2} I_q \sin 2\omega t + \sum_{n=2}^{\infty} \sqrt{2} I_{2n-1} \{\cos[2(n-1)\omega t + \varphi_{2n-1}] + \cos(2n\omega t + \varphi_{2n-1})\}$$

Likewise, the type (7) nearby two with rides $2\cos\omega t$:

$$2i_s(t)\cos\omega t = \sqrt{2}I_q + \sqrt{2}I_q\cos 2\omega t + \sqrt{2}I_p\sin 2\omega t + \\ + \sum_{n=2}^{\infty}\sqrt{2}I_{2n-1}\{\sin(2n\omega t - \varphi_{2n-1}) + \sin[2(n-1)\omega t + \varphi_{2n-1}]\} \tag{9}$$

Based on the wattful current separation's overtone examination control diagram like Figure 3.1, according to the above analysis, the electrical network electric current expands two times, multiplied by $\sin\omega t$, is lower than two time of electric current base frequency after the cut-off frequency low pass filter (LPF) may obtain $\sqrt{2}I_p$ [5], multiplied by $\sin\omega t$ obtains i_p ; Likewise may obtain i_q . This may act according to the user the choice, simultaneously carries on the compensation to the overtone and the idle current or only carries on the compensation to the overtone. If simultaneously carries on the compensation to the overtone and the idle current, may separate i_q channels, or adopts $i_p + i_h$.

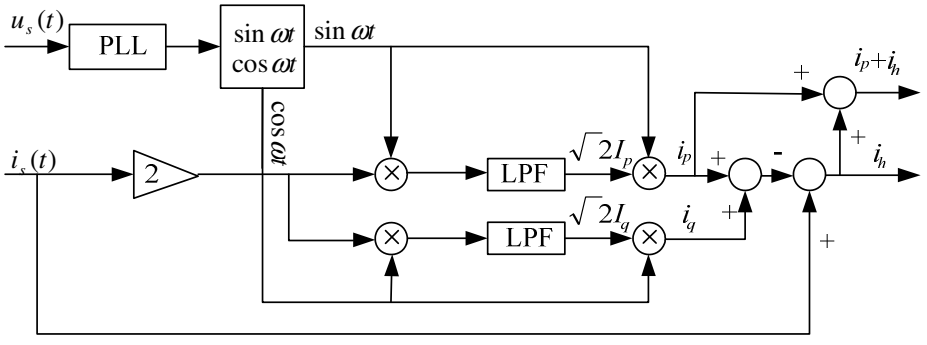


Fig. 3.1. Wattful current separation overtone examination control diagram

4 Specific Power Factor Overtone and Idle Current Examination Law

Supposes the network voltage expression same type (1), after supposing the compensation the electrical network electric current is:

$$i_s'(t) = ku_s(t) = k\sqrt{2}U\sin\omega t \tag{10}$$

Before supposing the compensation, electrical network electric current Fourier expansion is:

$$i_s(t) = \sum_{n=1}^{\infty} i_n\sin(n\omega t + \varphi_n) = ku_s(t) + i_q(t) + i_h(t) \tag{11}$$

Above equation 1 does not have the active power, namely satisfies:

$$\frac{1}{T}\int_{t=0}^{T} u_s(t)[i_q(t)+i_h(t)]dt = 0 \tag{12}$$

(10) substitution type (11):

$$k = \frac{\frac{1}{T}\int_{t=0}^{T} u_s(t)i_s(t)dt}{\frac{1}{T}\int_{t=0}^{T} u_s^2(t)dt} = \frac{\overline{u_s(t)i_s(t)}}{\overline{u_s^2(t)}} \tag{13}$$

We can obtain the improvement by the above analysis based on the specific power factor harmonic current examination method, like Figure 4.1. This method directly examine electrical network's idle work and the harmonic component by the network voltage and the electric current.

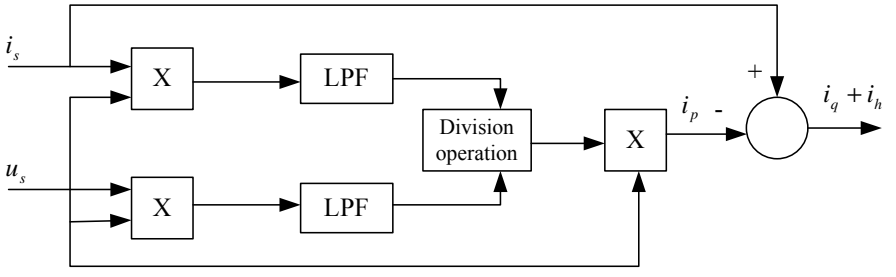


Fig. 4.1. Improvement based on specific power factor harmonic current examination diagram

5 MATLAB Simulation

Selection for the electricity iron single-phase APF overtone idle current examination method, applied in MATLAB7.0 simulink to analyze three kind of single-phase APF.

The supposition electricity iron tows arm's load voltage (the unit: V), electric current (unit: A) may use the type (14), (15) to simulate[6]:

$$u_s(t) = 27500\sin\omega t + 500\sin(3\omega t + 30°) \tag{14}$$

$$\begin{aligned} i_s(t) = 490[&\sin(\omega t + 30°) + 0.22\sin(3\omega t + 40°) + 0.10\sin(5\omega t + 18°) + \\ &0.07\sin(7\omega t + 30°) + 0.05\sin(9\omega t + 60°) + 0.02(11\omega t + 46°) + \\ &0.01\sin(13\omega t + 83°) + 0.01\sin 15\omega t + 0.01\sin(17\omega t + 138°) + \\ &0.01\sin(19\omega t + 20°) + 0.01\sin(21\omega t + 138°) + 0.01\sin 23\omega t] \end{aligned} \tag{15}$$

We may obtain the simulation result by the type :

Carries on the comparison to three kind of examination method's simulation result, gets the following conclusion:

(a)the idle current based on the instant reactive power theory examination law can only examine sum of the overtone and the idle current based on the specific power factor's overtone examination law;

(b)Three kind of examination's method structures is getting more and more simple, the examination precision to be getting more and more low, timeliness to be getting better and better.

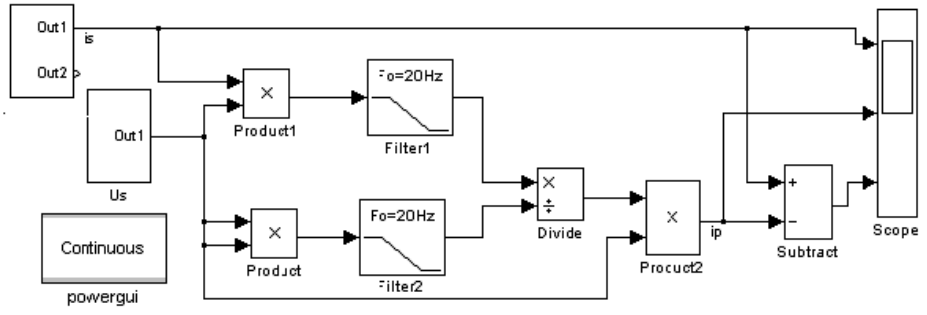


Fig. 5.1. Based on specific power factor overtone examination simulation model

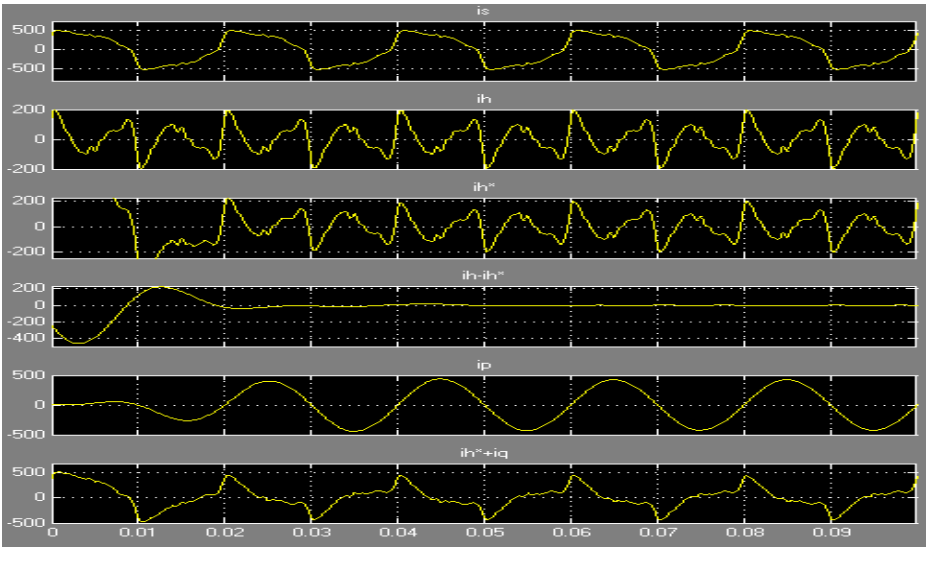


Fig 5.2. Instant reactive power theory examination method simulation result

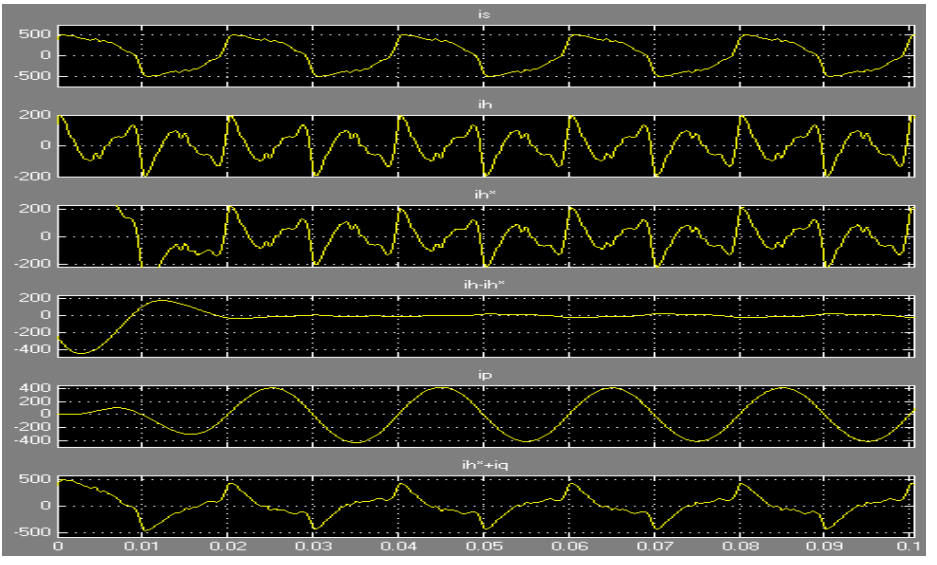


Fig. 5.3. Wattful current separation overtone examination simulation result

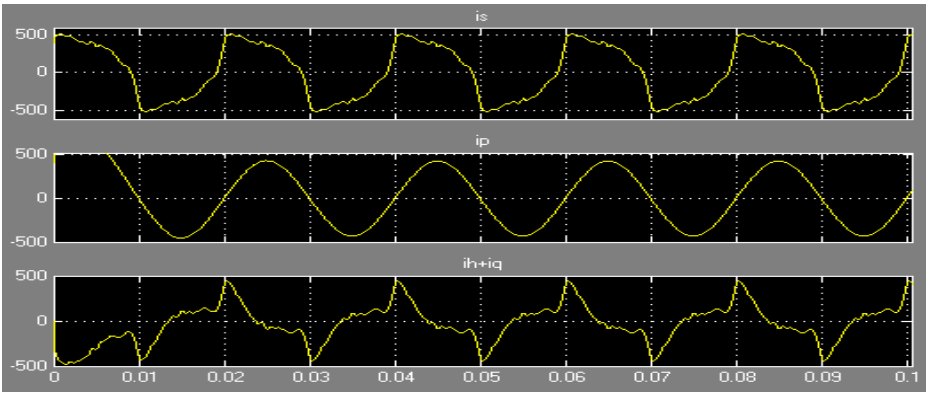


Fig. 5.4. Specific power factor overtone examination simulation result

6 Conclusions

From overall evaluation, take the electrification railroad hauling power supply system as compensating the object, uses the wattful current separation examination law effect is best. Therefore, this paper uses the overtone and idle current's examination based on the wattful current separation examination method.

References

1. Xie, G.R. (ed.): Power system grounding technology. Water Conservancy and Electric Power Press, Beijing (1991)
2. High-Voltage Power System, Tsinghua University compiled technical expertise. In: The Impact of High-Current Technology. Science Press, Beijing (September 1978)
3. Shen, L.S.: High-Speed Data Acquisition System Theory and Application. People's Posts & Telecom Press, Beijing (1995)
4. Liu, H.P.: TMS320LF240X DSP C language development and application. Beijing University of Aeronautics and Astronautics Press, Beijing (2003)
5. Liu, H.P.: TMS320LF240X DSP structure, principle and application. Beijing University of Aeronautics and Astronautics Press, Beijing (2002)
6. Chen, Y.Y.: MATLAB signal processing Xiang Jie. People's Posts & Telecom Press, Beijing (2001)

Human Action Recognition Based on Key Frames

Yong Hu and Wei Zheng

Jinling Institute of Technology
99 Hongjing Avenue, Nanjing 211169, China
{huyong,zhengwei}@jit.edu.cn

Abstract. Human action recognition has been an active research area over two decades. Although lots of approaches have been proposed in the literature with varying degrees of success, most of the approaches require lots of parameter estimation and computation time. In this study, we show that, human actions can be simply represented by key frames of the action sequence. The advantage of the approach is that the less the number of actions required and the features dimension is reduced. In addition, computational cost is decreased. We perform action recognition experiments on the Weizmann database, using features extracted from key frames. Experiments show that the proposed approach is effective and simple to implement.

Keywords: Human action recognition, key frames, polynomial fit, Weizmann database.

1 Introduction

Human action recognition has been an active research area over two decades. The objective of this research is to recognize human actions from videos so that the system can understand human behaviors and make further semantic description of the scene [1]. It has a wide range of applications such as visual surveillance, human-computer interfaces, digital entertainment and security system.

Vision-based human action recognition can be regarded as a combination of feature extraction, and subsequent classification of these image representations. There are several existing surveys within the area of vision-based human motion analysis and recognition [2]. In general, action recognition approaches can be categorized into one of the three groups: (i) template matching, (ii) state-space approaches and (iii) semantic description of human behaviors. Approaches based on template matching first convert an image sequence into a static shape pattern, and then compare it to pre-stored action prototypes during recognition [3]. The state-space approaches define each static posture as a state, and use certain probabilities to generate mutual connections between these states. Motion sequence was considered as a tour through the states of static postures. Through these tours, joint probability with the maximum value is selected as the criterion for action classification. The purpose of semantic description of human behaviors is choosing a group of motion words to represent the human behaviors in natural scenes [4].

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 535–542, 2011.

Although lots of approaches have been proposed and developed in the literature with varying degrees of success, most of them analyze an entire video sequences. These approaches require lots of parameter estimation and result in large computational cost and computation time. In this paper, we present a scheme for action recognition based on key frames. The advantage of the approach is that the less the number of actions required and the features dimension is reduced. In addition, computational cost is decreased. We prove that, human actions can be simply represented by key frames in one cycle of the whole action sequence. We perform action recognition experiments on the Weizmann database, and experiments show that the proposed approach is effective and easy to implement.

This paper is organized as follows. In Section 2, we briefly introduce the materials and methods used in this paper. In Section 3, the proposed scheme for human action recognition, includes automatic extraction of key frames and feature extraction, is presented in detail. Experimental results are reported in section 4. Finally, the conclusion is provided in Section 5.

2 Materials and Methods

In this section, we briefly introduce the materials and methods used in this paper: Zernike moments, shape descriptors and SVM classifier.

2.1 Zernike Moments and Shape Descriptors

Zernike moments have been widely used in human action recognition as holistic features. It has the characteristics of rotation invariant and robust to noise. Scale and translation invariance can be implemented using normalization. The Zernike moments of image intensity ***f(x,y)*** with order ***n*** and repetition ***m*** is expressed as follows.

$$Z_{nm} = \frac{n+1}{\lambda} \sum_{x} \sum_{y} f(x, y) V^{*}_{nm}(x, y) \tag{1}$$

where λ is a normalization factor and $\boldsymbol{x}^2+\boldsymbol{y}^2 \leq 1$.

To achieve scale and translation invariance, the image function ***f(x,y,t)*** can be normalized by transforming it into ***g(x,y,t)***. The transforming equation is list below.

$$g(x, y, t) = f(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y}) \tag{2}$$

Compactness is the measurement of describing the regional compact of an object. It is often defined as the ratio of the squared perimeter and the area. It reaches the minimum in a circular object and approaches infinity in thin, complex objects. The formula is as follows.

$$Comp = \frac{2\sqrt{A\pi}}{P} \tag{3}$$

where ***A*** is the area, and ***P*** is the perimeter of the object.

In this paper, the magnitude of Zernike moments was combined with shape descriptors as descriptive features. All features were computed from one action cycle, and then a polynomial ***p(x)*** of degree ***n*** is fitted to the values of the features.

2.2 SVM Classifier

Support vector machine (SVM) is an effective classification technique on the foundation of statistical learning theory. In the SVM method, a unique principal called structural risk minimization (SRM) principle is used to minimize an upper bound of the generalization error. Another key point of the SVM is that training SVM is equivalent to solving a linearly constrained quadratic programming problem so that the solution of the SVM is always unique and globally optimal.

Consider the two-class linear separable training sets

$$(x_i, y_i), \quad i = 1, \cdots, n, \quad x \in R^d, \quad y \in \{-1, +1\} \tag{4}$$

if there is a hyperplane

$$\omega \bullet x + b = 0 \tag{5}$$

all samples can be correctly classified without error and the separation margin can be maximized, then the hyperplane is so-called optimal hyperplane.

The optimal separating hyperplane is computed as a decision surface of the form:

$$f(x) = \operatorname{sgn}\{\sum_{i=1}^{m} \alpha_i^* y_i (x_i \bullet x) + b^*\} \tag{6}$$

In this study, we perform LIBSVM for multi-class classification. The LIBSVM is a library for support vector machines which is available in [5]. By using the software, a cross-validation method was utilized to analyze the classification performance. All labeled samples were split into two subsets: 8 subjects are randomly selected for training and others for testing. The trained SVM classifier predicts the action label for the testing actions. This procedure was repeated 5 times, and the average results are listed in Table 1.

3 The Proposed Framework

Human motion analysis concerns on three major issues, detection, tracking and recognition from image sequences. The recognition of movement can be performed at various levels of abstraction. The common taxonomy is composed of three categories: action primitive, action and activity. An action primitive is an atomic movement that can be described at the limb level. An action consists of action primitives and describes a whole-body movement. An activity contains a number of subsequent actions, and gives an interpretation of the movement that is being performed [2].

In this study, we only focus on the most important part of human motion analysis, action recognition. There are two parts in action recognition approaches, image representation and action classification. Some of the image representations extract image features for each frame in the sequence individually. Unlike other existed approaches, we extract features only from one cycle. In this part, key frames were automatic extracted from the image sequence, and one action cycle was extracted from the whole sequence. And then, features were extracted by Zernike moments and shape descriptors. A polynomial ***p(x)*** of degree n is fitted to the values of the features of each

frame. In the action classification part, a SVM classifier was used for training and testing. The resulting approach was applied to the recognition of human actions and the performance was compared to other classic approaches.

3.1 Automatic Extraction of Key Frames

In the area of human action recognition, few approaches were proposed based on key frames. Schindler [6] present a system for action recognition from very short sequences ("snippets") of 1-10 frames, and systematically evaluate it on standard datasets. Recognition results show that short action snippets are almost as informative as the entire video. Lv [7] presents an approach that does not explicitly infer 3D pose at each frame. The proposed approach was demonstrated on the Inria Xmas dataset and achieved an overall action recognition rate of 80.6% using a single camera. Carlsson [8] consider the recognition of a specific action (forehand and backhand strokes) of a specific person. They present an algorithm for localizing action specific frames in video sequences. Their experiments show that using key-frames within the sequence all forehand strokes are found with no false positives.

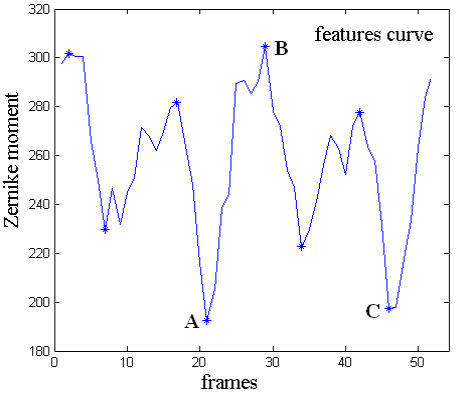


Fig. 1. Feature curve of "denis_jack" in Weizmann database

Given an image sequence of human action, theoretically all frames in the sequence should be exploited for recognition of the action. However, human vision proves that frequently human actions can be recognized from few frames almost instantaneously. Figure 1 shows a feature curve of an action "jack" in Weizmann database [9], each value on this curve is according to the Zernike moment of the silhouette. From figure 1, we can see that there are some local maximum and some local minimum points on this curve (marked with '*'). These points imply motion corners, indicating the person's motion changes at these positions. Frames corresponding to these extrema are called action key-frames (See Figure 2). Each segment between two minimum points (such as A to C) or maximum points is one cycle of a repetitive motion.

Fig. 2. Frames corresponding to extrema points

Using these key frames, a compact representation of a human action is obtained for action recognition. Extract from one cycle of the whole action sequence is more efficient and the computational cost is decreased.

3.2 Feature Extraction

In this study, we propose a different way of exploiting the information contained in the motion cycle. First, holistic features were computed based on each single frame in one cycle. The features were extracted by Zernike moments and shape descriptors. And then, a polynomial ***p(x)*** of degree n is fitted to the values of the features of each frame (See Figure 3). Finally, the polynomials coefficients were constructed to a new vector as descriptive features.

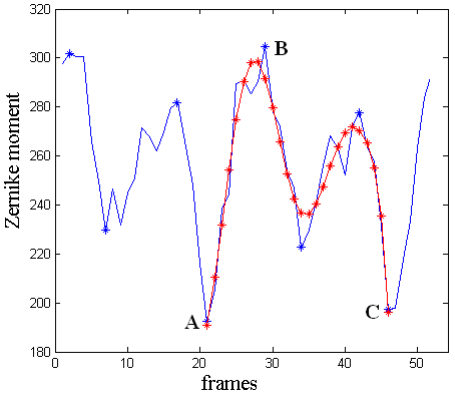


Fig. 3. An eighth-order polynomial fit of the given curve

The polynomial is defined as follows.

$$p(x) = p_1 x^n + p_2 x^{n-1} + \cdots\cdots + p_n x + p_{n+1} \tag{7}$$

The new features describe the tendencies of variation of each frame. Experimental investigations indicate that the polynomials coefficients are more robust than the features themselves.

4 Experimental Results

To evaluate the proposed method, we perform our experiment on the Weizmann dataset [9], which is recorded at the Weizmann institute. It was consists of 10 different actions: bend, gallop sideways, jump, jump in place, jumping jack, one- hand wave, run, skip, two-hand wave and walk. Each action was performed by 10 persons at varying speeds. The viewpoint and the backgrounds are static; foreground silhouettes are included in the dataset. The dataset have been widely used to evaluate action recognition approaches and many results have been reported. Figure 1 shows some example frames of this dataset.

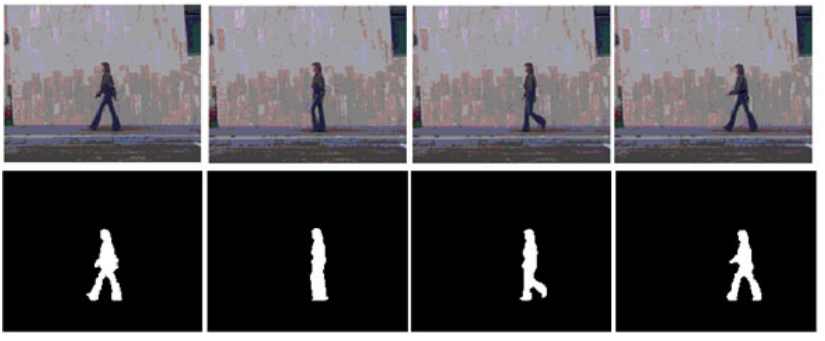


Fig. 4. Weizmann dataset

Tables 1 represent the results of three different methods. They share the same framework except features extraction. The first one extracted features from each frame in whole sequence while the second on extracted features only from one cycle. Numbers in bold print denote the best results of the five times.

Table 1. Classification results

Frames	Features	Recognition Rate					
		5 times					Aver.
each frame in whole sequence	Zernike moments & shape descriptor	87.1%	87.9%	90.6%	**92.0%**	89.8%	89.5%
each frame in one cycle	Zernike moments & shape descriptor	86.9%	**92.8%**	90.1%	88.5%	88.1%	89.3%
each frame in one cycle	Polynomials coefficients	92.4%	91.0%	**94.8%**	92.6%	93.8%	92.9%

It shows in the table that the proposed method outperforms the others. The recognition rate for our proposed method is 92.9%, while the others are 89.5% and 89.3%, respectively. Overall, our method shows higher efficiency and accuracy than the comparison algorithm. This shows that the polynomials coefficients achieved better performance than features themselves.

The confusion matrix of our method is shown in Figure 4. It can be seen that some videos are partially misclassified. Most of the mistakes are intuitively reasonable, e.g., “one-hand wave” is confused with “two-hand wave”. The most confusion occurs between run, walk and gallop sideways. Run, walk and gallop sideways are often misrecognized since they are quite similar actions in lower body parts.

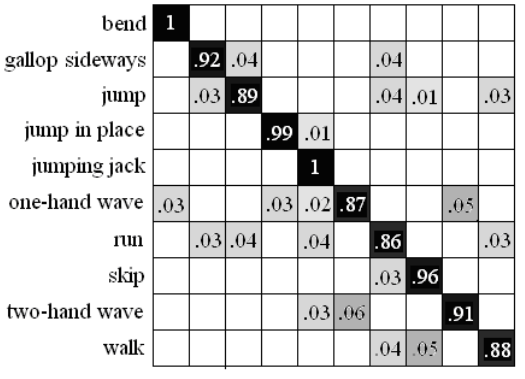


Fig. 5. Confusion matrix

5 Summary and Conclusions

In this study, we proposed a human action recognition approach based on key frames. We show that, human actions can be simply represented by key frames of the action sequence. The features were extracted in each frame by using Zernike moments and shape descriptors. And then, a polynomial is fitted to the values of the features. The polynomials coefficients were considered as descriptive features for training a SVM classifier. We perform the approach on the Weizmann database, and experiments show that the proposed approach is effective and simple to implement.

Acknowledgments. This work is supported by Research Foundation of Jinling Institute of Technology (Grant No. 40410827).

References

1. Liu, C., Yuen, P.C.: Human action recognition using boosted EigenActions. Image and Vision Computing 28(5), 825–835 (2010)
2. Poppe, R.: A survey on vision-based human action recognition. Image and Vision Computing 28(6), 976–990 (2010)
3. Ahad, M., Tan, J.K., Kim, H., Ishikawa, S.: Motion history image: its variants and applications. Machine Vision and Applications 21(6), 1–27 (2010)
4. Wang, L., Hu, W., Tan, T.: Recent developments in human motion analysis. Pattern Recognition 36(3), 585–601 (2003)

5. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. Software (2001), http://www.csie.ntu.edu.tw/~cjlin/libsvm
6. Schindler, K., Gool, L.V.: Action Snippets: How many frames does human action recognition require? In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008), pp. 1–8. IEEE Press, Anchorage (2008)
7. Lv, F., Nevatia, R.: Single View Human Action Recognition using Key Pose Matching and Viterbi Path Searching. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007), pp. 1–8. IEEE Press, Minneapolis (2007)
8. Carlsson, S., Sullivan, J.: Action Recognition by Shape Matching to Key Frames. In: IEEE Computer Society Workshop on Models Versus Exemplars in Computer Vision, pp. 1–8. IEEE Press, Hawaii (2001)
9. Blank, M., Gorelick, L., Shechtman, E., et al.: Actions as space-time shapes. In: 10th IEEE International Conference on Computer Vision (ICCV 2005), pp. 1395–1402. IEEE Press, Beijing (2005)

A Novel Interactive Device for 3D Display

Wankui Liu[1], Yuan Fu[2], Yi Yang[3], Zhonghong Shen[3], and Yue Liu[1]

[1] School of Optics and Electronics, Beijing Institute of Technology, Beijing, China
[2] The High School Affiliated to Renmin University of China
[3] School of Automation, Beijing Institute of Technology, Beijing, China

Abstract. Though 3D displays become more and more prevalent nowadays, few of them have interactive devices. In this paper, a novel simple interactive 3D input device called touch box is introduced, which just consist of two touch resistive panels. Most of the 3D pose can be simulated by these touch panels. To show how the touch box works, a 3D game application is implemented and a set of 3D manipulations including translating, scaling, rotating and so on are listed. More future work is discussed at last.

Keywords: 3D display, interactions, input interface.

1 Introduction

3D displays offer a strong sense of immersion to users. However, usually there are few interactions in the virtual environments. Considering the price of the 3D control products available, a low cost and simple 3D control is needed. That's our main purpose.

In this paper, we present an implementation of a novel, multi-touch-sensitive 3D display. First, we list mainly related work on 3D touch control. Second, we describe hardware and software components needed to facilitate 3D touch control of the 3D displays. The principle of this system is also included. Third, we present a set of touch interaction techniques, mainly including the translation, scaling, rotating. The gestures are divided into two groups: one hand touch and two hands touch.

2 Related Work

Considering the normal 3D display, we want to have a 3D interactive application. Most 3D controls use trackballs by one hand. Data glove, tracker and joystick are also mainly used 3D input devices.

In [1], some different solution is proposed for large Displays: two infrared LEDs, hold-up tool card, foot-operated tool cards and infrared pen with a click button. 3Dconnexion SpaceNavigator[2] is a very impressive 3D mouse, which can give excellent 3D pose change function.

CubTile[3] is a multi-touch cubic interface, which supports lots of interaction methods. Five surfaces can finish some 3D interaction task perfectly. There are also

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 543–548, 2011.

some studies on 3D interactions with multi-touch devices, like [4, 5], which navigate the 3D environment with just two or three fingers. Similarly two points input can be used to achieve 3D effects [6].

Considering the price and the complexity, we present our touch box just with two touch panels each side and list a set of concept of interactions. With 3D mouse, the 3D touch sense can make the virtual environment more real.

3 System Overview

3.1 Hardware Setup

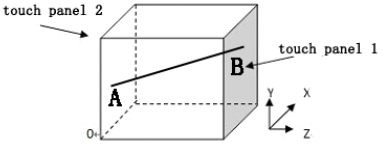


Fig. 1. Touch box

The touch box consists of two resistive touch panels, as Fig.1 shows. When the user touches one panel, the position information can be calculated and transferred to the same side processor. Then the side processor calculates the coordinates and sends the coordinates to the center processor. With another side processor data, the center processor figures out the angles between the line AB (point A and point B are all touch points) and each side of the touch box. With this message, we can achieve a set of interaction, including moving, scaling, rotating, changing direction and so on.

The inspiration for the 3D display comes from the observation of the LCD with polarized glass. Mirror just reflects the polarized light changing the polarization direction perpendicularly, while mirror refracts the polarized light unchanging the polarization direction. Therefore, a 3D display is made [7], which is a little like StereoMirror 3D display [8] but a group of adjustment is made to achieve better performance. This 3D display system consists of two LCD monitors in an "open book" configuration with a bisecting (45°) half-silvered mirror. The study process is as Fig.2 shows.

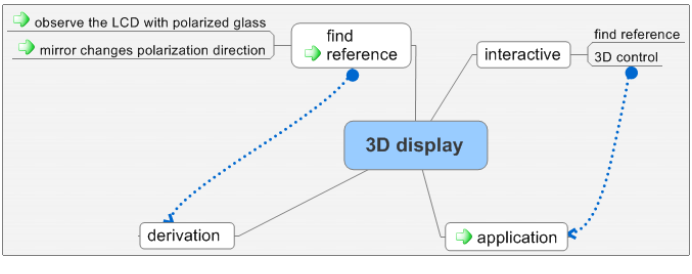


Fig. 2. The process of this 3D display 3D control study

3.2 Control System

The touch box contain two parts: one part is data acquisition module and the other one is data transmission interface module part.

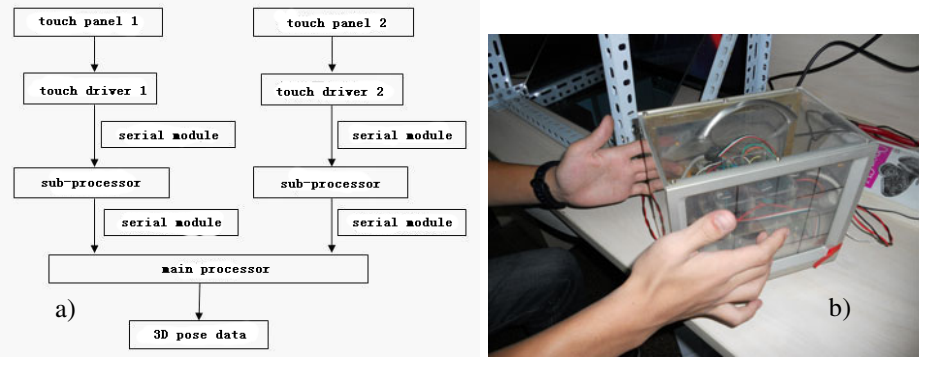


Fig. 3. (a)The structure of the touch box (b) the touch box

The touch panel offers the touch information to the sub-processor, which will calculate the position of the touch point. After combined the data of the two side sub-processors, the main processor present a set of data including the position of the two touch point, the angle between the line AB(Fig.1(b)) and three different direction edges. The touch box can send out touch data by 9600kbps, 9 bytes per frame with a start flag byte and an end flag byte. Owing to these data, we can obtain the 3D pose change expediently.

As shown in figure 4, we adopt Atmega 128 single chip produced by AVR company as the system processor in the control system. We use a touch panel of five line type resistance as the touch sensor to sense the touch point, and the processor communicate with touch panel driver via RS232 serial communication agreement. A power supply module developed by ourselves (7805 power chip as the main components) is employed in this system, which produce a 5v DC and a 12v DC. The former supplys power for single chips, while the latter affords power for the touch panels and drivers. For the touch-panel driver sends location information of touch points outside uniaxially by RS232 serial port, and one touch panel driver only can control one touch panel, this system takes use of three single chip processors, two of them (slave processor) are used to receive and resolve the location information of touch points, while the other one (main processor) is used to analyse the location information of touch points and conclude the 3D pose information of the line between the two touch points. A cube named touch box consists of processor module, power supply module, touch panel driver module, touch panel senor module, and communicate with outside via RS232 Serial port.

In the software control of this system, for the touch panel driver sends data out uniaxially, we utilize a much information contrast calibration way to insure the accuracy of data which comes from the touch panel, that is, we compare the data comes from the touch panel driver in a certain time (millisecond timescale) and

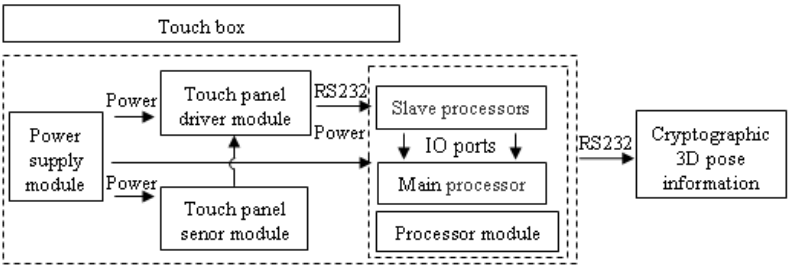


Fig. 4. The hardware system of the touch box

analyse their mean and variance, if the two value are both in a certain accuracy range, then we believe that the data is correct reasonably and can use it in the next part. The program of main processor calculates and analyses the 3D pose information of the line shaped between two touched points mainly, and issue control information. Hence, the running cycle of main program (T) is short enough so that it can meet the system requirement of real-time, exactly, the two slave processors send location information of touched points via IO ports and the main processor receives these information from the two slave processors in the light of a certain time interval which approximates as T. Besides, the main processor encrypt the 3D pose information of the line shaped between two touched points to insure this information can be received and used precisely by subsequent module.

3.3 Game Application

To show how the touch box works, a 3D game called Spaceship is designed with Microsoft XNA Game studio 3.1. The touch box can be used to control the spaceship, changing its position and avioding or firing at the meteors. The game first checks the touch data, filtering noises and false information. From the exact touch data, the gesture can be judged and then according to the gesture definition, the game will correspond to attitude correction of the spaceship, including translating, scaling, rotating.

The game at last renders the all scenes to two textures. The scenes are rendered by two different cameras with different view angles. One simulates left eye and the other one simulates the right eye. They are output to the top LCD and the back LCD respectively. Then a 3D game is formed. The game can draw these textures with 27 frames per second.

4 3D Manipulation

In this game application, there are some touch gestures we defined. They are as Follows.

1. One side touch. Since each side of the touch box is flat and 2D, it will support 2 degrees of freedom. A single touch on one side of the touch box can make the 3D model move forward, backward, upward or downward by moving the touching finger.

2. Two sides touch with the same moving trend. Users can scale the 3D model by moving all their fingers upward or downward. When the two touch points are all rotating in the same speed and direction with the same radius, the 3D model can fulfill pitch motion. The 3D model will move far or near the panel if the user moves both his fingers forward or backward.
3. Two sides touch with the different moving trend. The roll and yaw motion can be operated using two fingers moving at the opposite direction horizontally and vertically.

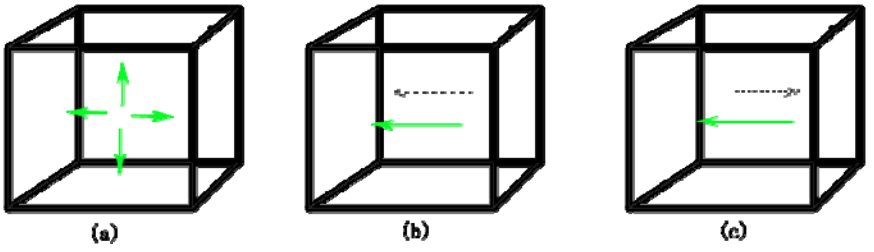


Fig. 5. Interactions (a)one side touch (b)two sides touch with the same moving trend (c)two sides touch with the different moving trend

These interactions are enough for a single user to fulfill six-degree-of-freedom motion, especially for our game. As the touch box sends the serials bytes information, transporting the position and the 3D pose, the users can deal with this information by themselves. The definition of our touch box interaction gesture is not limited. Therefore, the users can define the touch gesture by themselves as they need. Therefore this touch box can be treated as a 3D mouse.

This touch box is suitable for navigation tasks, which can perfectly achieve translating, scaling, rotating and so on effects. Compared to CubTile, only two touch panels are used expediently. The next generation touch box will be improved on the resolution, accuracy and stability. The gesture will be defined in the touch box main processor, and more novel concept of the touch gesture will be integrated.

5 Conclusion and Future Work

We have presented a novel interactive 3D control. The hardware and software solution are all presented. The interaction with two touch panel to simulate 3D motion is introduced. With the touch box, the user can be enabled to achieve 3D tour effect. Meanwhile, the interface is open and the user can define the touch gesture themselves. In the future, much research remains to be done. This touch box will be compact to a smaller one. Besides, the touch panel may be changed to better one, of which the resolution and accuracy should be what we expect. We also want to explore the map function so that the user can look through Google Earth freely by the touch box. More interestingly, we want enable this touch box control the windows system in 3D method, such as navigating the desktop. There are also other great novelties, which make that we will continue to meet interesting research challenges.

Acknowledgement

This work was supported by Program for Changjiang Scholars and Innovative Research Team in University of Ministry of Education of China.

References

1. Laundry, B., Zealand, N.: Interaction with 3D Models on Large Displays using 3D Input Techniques. In: CHINZ 2010 Proceedings of the 11th International Conference of the NZ Chapter of the ACM Special Interest Group on Human-Computer Interaction. ACM, New York (2010) ISBN: 978-1-4503-0104-6
2. 3Dconnexion, http://www.3dconnexion.com
3. dl Riviere, J.-B., Kervegant, C., Orvain, E., Dittlo, N.: CubTile: A multi-touch cubic interface. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST, Bordeaux, pp. 69–72 (2008)
4. Cardinaels, M., Frederix, K., Nulens, J., van Rijsselbergen, D., Verwaest, M., Bekaert, P.: A multi-touch 3D set modeler for drama production. Digital Media, 330–335 (2008)
5. Edelmann, J., Fleck, S., Schilling, A.: The DabR– A Multitouch System for Intuitive 3D Scene Navigation. In: 3DTV CON the True Vision, pp. 1–4 (2009)
6. Zeleznik, R.C., Forsberg, A.S., Strauss, P.S.: Two pointer input for 3D interaction. In: Proceedings of the 1997 Symposium on Interactive 3D graphics, pp. 115–120. ACM Press, New York (1997)
7. Liu, W., Liu, Y.: A novel interactive 3D display based on half silvered mirror. In: ICMEAT 2011 (2011)
8. Hochbaum, A., Fergason, J.L.: Optical design considerations for a beam combiner in a StereoMirror™ 3D display. SPIE, vol. 6055, 60551I (2006)

Design and Application of Hybrid Learning Platform Based on Joomla

Li-sen Qi and Ai-kui Tian

School of Computer Science and Technology, Shandong University of Technology, 255049 Zibo, Shandong, China

Abstract. Joomla is a typical representative of open source CMS, which has great application value and development potential in the adult teaching. A hybrid learning platform can be quickly erected based on Joomla, and supports management of learning content, learning methods, learning resources, e-learning files, video-on-demand and other aspects, and presents a line of thought for open curriculum construction.

Keywords: hybrid learning, Joomla, adult education.

1 Introduction

Today, the development of information technology has been constantly improving in all walks of life. From the content management systems to web 2.0 as well as the semantic web, a variety of novel concepts are constantly coming to the fore, people's awareness of the world has been quietly undergoing changes. Blog, podcast technology as the representative, the internet application technology around interactive and personalized experience for users has been improved and come into mature. Due to the emergence of new internet technologies, people can be freed from the complex network technology, and more involve in the creative work of their own knowledge. At the same time, changes in the social environment lead to a more flexible and diversified ways of learning, the learning environment continues to optimize and innovate, the effect of learning continues to improve significantly.

As the effective mode of learning in the current education field, hybrid learning is attracting more and more concern among the workers of educational technology. It is the combination of traditional teaching methods and network teaching methods, which reflects "teacher initiative—student oriented" teaching mode, and the optimizing results of teaching elements under information technology environment. In recent years, hybrid learning theory research and practice cases are beginning to show the trend of popularity, for example, Gao-feng Ruan and Xiao-dong Xu (2009) introduced constructing a blended learning environment based on public internet services based on their course of “Human Learning and Cognition”, Ling Qian and Wu-lin Ma (2009) discussed how to expand hybrid teaching of college English curriculum based on web 2.0. Wei Yi (2005) argued that the hybrid learning was a mixture of learning methods, learning media, learning content, learning models and student support services and learning environment, and built an integrated learning system architecture to support hybrid learning model.[1] In addition, it is more common for many Chinese scholars to

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 549–556, 2011.

build their learning platform through the Blog, Moodle, Google Site and other systems, and the cases have been a certain degree of promotion in the teaching practice of adult education.

2 Joomla's Main Features and Technical Framework

Currently, Joomla, Drupal, PHPcms are all the mainstream CMS within the world, their user groups are relatively large. As an open-source CMS system based on Apache + php + mysql model, Joomla occupies the mainstream position in the business website system with prominent features of a powerful plug-in expansion, the human back-end management, and variety of template styles, and so on. Its biggest feature is easy to administer, can organize units, classifications, articles, templates under unified management. In addition, the open source characteristics of Joomla make it ease to install in a Windows, Linux, MacOSX and other systems platform, and can be carried out by installing the plug-in extensions, as shown in Fig.1.

Fig. 1. The control panel of Joomla back-end system

2.1 Menu Management

The menu management of Joomla includes the main menu, user menu, top menu, resources, sample pages, key concepts and some others. The navigation menu is the most important part of the entire site layout, and it has an important influence on the user's access behavior and website operations. Fully awaring of the importance of navigation menu helps achieve the design uniqueness and scientific layout.

2.2 Content Management

Content Management is one of the characteristic functions of Joomla, which specifically divides into unit management, category management and articles management, and

among them is mutually subordinate relationship, in other words, the article belongs to a classification, the classification belongs to a unit. In general, the articles can not exist in isolation, and must belong to a particular unit and its classification. When you create a menu, you can set the menu item type, which will decide the way that the article shows. You can display them in an article, display all articles under a category, and also display the category list under a unit.

2.3 Component Management

Joomla itself provides a system framework, in which the menu, login, advertising, voting, useful links are all organized in accordance with module approach. Administrator may install and uninstall specific modules at any time, such as using the JCE Administrator to visualize the article editing, using the Easy Guestbook for user messages, using the SWMenuFree for production of dynamic menus, etc.

2.4 Template Management

Joomla 1.5 realized the MVC architecture, HTML, CSS, etc. and the relevant display parts are separated from the logic and data part, all allows us to easily control the appearance, without concerning how the core code operates. The back-end of Joomla provides template management functionality which can be easily to switch templates among different styles, and can be customized according to individual needs. Currently, the template resources are very rich in the Internet, not only have charging templates represented by JoomlArt, but also free templates shared by users.

3 Designing Hybrid Learning Platform Based on Joomla

In general, the characteristics of traditional instructional design are teaching oriented; and the modern teaching design is learning oriented mostly. From the theoretical foundation and practical design perspective, hybrid learning is the combination of “teaching-based” instruction design and “learning-based” instruction design. Through the hybrid learning concept and design, we have developed a "C language" online course, which achieved a good result after a year of teaching implementation. "C language" teaching at school of continuing education, Shandong University of Technology, has employed open source technologies of Joomla, combined with daily teaching practice, and established "C Language" online course to “confront the adult education, serve the public” and turns out to be a typical example of hybrid teaching model.

As a professional programming language, C language has powerful function, high efficiency, practicability and so on. On the other hand, there are many shortcomings when C-language is chosen as a universal-type programming entry language. We should think about how to take effective measures in teaching, so as to stay away from obscure syntax for beginners, to start learning C programming language from the basic, the most practical way, and strive to enable students to learn the basic programming method as soon as possible, proceeding to reach the realm of applying programming to solve practical problems.

From the learners point of view,the purpose of learning C language is to learn to write programs to solve practical problems, rather than become skillfully proficient in C language syntax, and grasp various programming skills. In response to these problems, we propose a new teaching method - application of demand-driven approach to the C language teaching. We assign C language boring and convoluted syntax of knowledge to the various chapters, namely, in accordance with the demands of programming application, gradually introducing the corresponding grammar knowledge and programming methods. Based on the above understanding, we design the hybrid learning process,as shown in Fig.2.

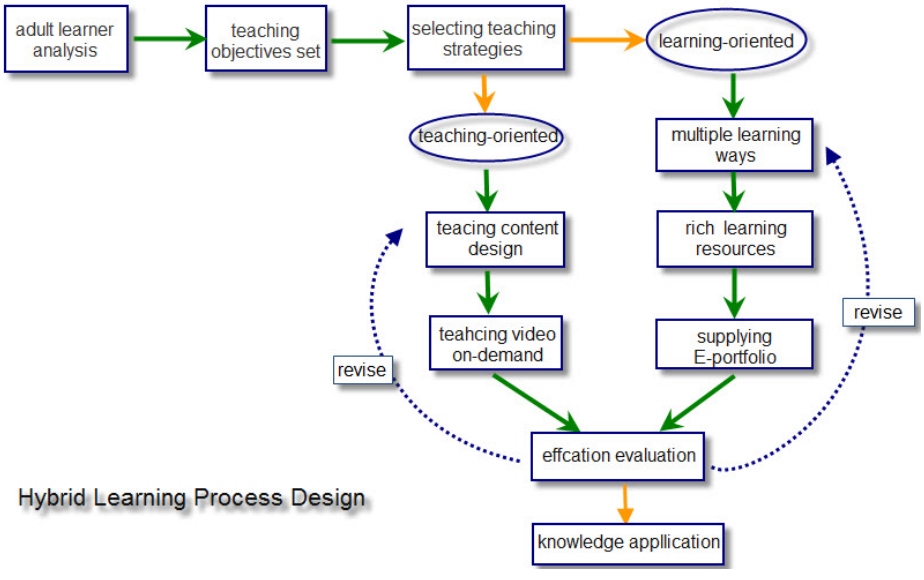


Fig. 2. The design of hybrid learning process

3.1 Teaching Content Design

In accordance with instructional design of hybrid learning, guiding learning-support services, adult education learning platform should provide a large number of learning resources to meet the learning needs of students at various levels. To achieve this end, we should carry out a detailed design to the organizational structure of the curriculum website, not only including graphics-based curriculum knowledge, but also video-based instruction recording; not only including the knowledge elements of the course, but also including a detailed self-study guide; including not only E-learning portfolio submission, but also the launching of student-teacher interaction. The overall design of "C language" online course is shown in Fig.3.

3.2 Teaching Video on Demand

Due to regional differences in network environment, the learning progress of each student might be different. Online learning training videos need to be smoothly played

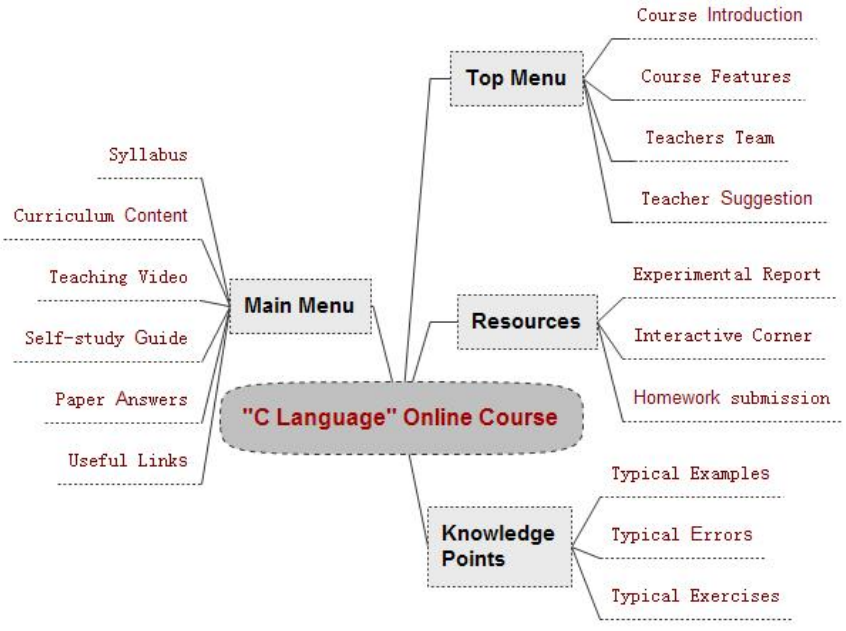


Fig. 3. The design of "C language" online course

under low-bandwidth environment, also random jumping according to chapters tips. Using streaming media technology, we suggest that video recording adopts 225Kbps bit rate, 320*240 resolution of the WMV format, and courseware adopts FLASH format. Streaming media can be understood as the use of streaming media technology to transmit multimedia files. Streaming media is a kind of transmission rather than a media in fact. Streaming media technology enables to play video when the cache size is not too large, just in a few seconds delay, and you can drag to watch the clear video. It is an effective solution to the shortcomings of downloading all the files needed from the internet to play, and it saves minutes or even hours of time for low bandwidth users.

We take into account many factors such as application environment, and use Windows Media Services component integrated into Microsoft Windows2003 to build streaming media server. The components need to install a Windows component in the Add or Remove Programs option of Windows 2003. After the installation has finished, you can use this component to create video release point. The name of the release point is called the name of streaming media, and it can be a single video file, multiple video files of a directory or live streams. After the recorded videos are released, the courseware which consists of three split-screen only to call this release name, and the whole process will be ended. When you succeed in streaming media playback, there will be a message "optimized for streaming experience, please click .....". One webpage of teaching video on demand is shown in Fig.4.

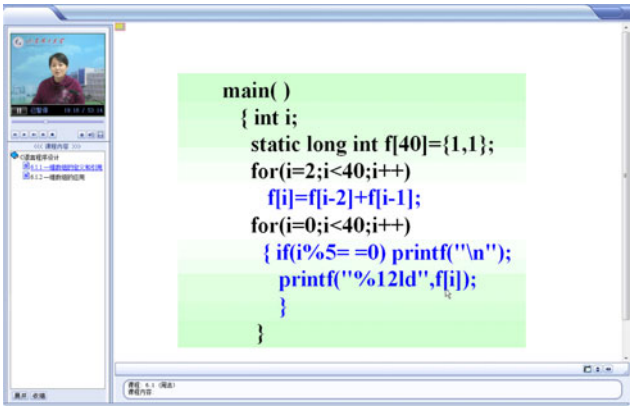


Fig. 4. The typical display interface of teaching video on demand: three split-screen

In hybrid learning, the arrangements of instruction elements are not static, and this affects the existing forms of interaction between teachers and students directly. The teacher-student relationship reflects the instruction theory deeply, the role of teachers and the status of students.Teachers are the teaching organizers and the facilitator of learning. Students are the absolute subject of hybrid learning, and only the initiative and enthusiasm of the students are stimulated fully so as to enhance the teaching effectiveness. In short, to achieve positive teacher-student relationship is the highest state which all educators pursuit.

3.3 Mutiple Teaching Organization Forms and Learning Ways

Instruction often occurs in very complex real environment, which requires to take flexible forms of teaching organization according to the actual need. In general, these forms are divided into three types: collective teaching, group instruction and individual instruction. The efficiency of collective teaching is the highest, followed by group instruction. Individual instruction is the lowest, but its effection is often best. Adult learners are constrainted by irregular study time, and collective teaching should not be adopted. Group instruction and individual instruction is more appropriate. Hybrid learning takes into account these factors fully, and increases the flexibility of teaching organization forms.

Since the hybrid learning has attracted academic attention, it was given the important task of integrating existing teaching methods. The complex nature of the educational environment determines there is no existence of a universal way of learning. Any way of learning has its specific advantages and disadvantages, and its value should be examined against its own historical background.

Hybrid learning has been considered as a mixture of collective learning and individual learning, a mixture of synchronous learning and asynchronous learning, a mixture of self-learning and collaborative learning, and also a mixture of formal learning and informal learning, which fully embodies the humanistic learning theory, and the value of life and education demands.

3.4 Rich Learning Resources

In adult education, textbooks, exercise books, and learning materials are very important for the implementation of teaching. Distance learning materials are different from the traditional material, greatly emphasize on instructional design, and reflect the communication effect of teaching media, and shift from paper books to video-audio tutorials, with characteristics of the vivid, intuitive and interesting teaching. An important feature of an e-learning system is the presentation way of the semantic relationships between topics using an appropriate navigational structure. [2]

Instruction resources are an important part of teaching and learning activities. For teachers, it includes well-structured course content, various forms of teaching material, and digital learning environment and cognitive tools. In hybrid learning based on Joomla, the teacher can apply instruction resources into the teaching process flexibly according to the actual instruction situation, For students, rich teaching resources increase the possibility of personalized learning, and support the knowledge transfer and innovation.

To improve the quality and effectiveness of online teaching video of the "C Language" course, we have organized dedicated curriculum development team, including online teachers, instructional designers, editors, to work together to record a high level of video courseware. In addition, for FAQs, we create introduction to knowledge and analytical elements, such as typical example, typical error and typical exercise, which has become an essential reference manual for student in their E-learning.

3.5 Supplying E-Portfolio

Lifelong learning within the frame of the information society appears to be inevitably related to the need for a flexible and just in time offer of education on issues which may be of highly focused interest. [3] To adapt to the development of adult education in information age, E-Portfolio system will gradually occupy the mainstream, and lead to change traditional learning methods and evaluation methods, to promote innovation of learning support services, and to promote the lifelong learning systems.

Hybrid learning evaluation is comprehensive assessment, value judgments and rational reflection of teaching process and results. For teachers, it re-examinates the instructional design process, provides a range of reference for future teaching adjustment. For learners, teaching evaluation can monitor learning process, fully reflect the learning performance and learning outcomes. From the perspective of developmental psychology, formative evaluation may be more true in the reflection of students status, but more complicated to implement than summative assessment. In order to make evaluation process more authentic and objective, a variety of evaluation methods should be advocated, such as assignment submission, E-Portfolio, online examinations,etc.

In China, the theoretical value and application of E-Portfolio has aroused the concern of scholars, but the mature application examples and system technology of E-Portfolio is still very lacking. The Joomla Flash Uploader (JFU) component under Joomla assists registered users to submit learning outcomes, such as test reports, daily operations, lesson reports, etc. The administrator can modify the parameters, in order to

assign different permissions to registered users, including the upload directory, and modifying the restrictions, which not only support the students to participate in the program, such as self-study, reflection, and make it easier to get teachers to formative assessment of students, and help shape a virtual learning community to share knowledge and wisdom.

4 Conclusions

The concepts of hybrid learning are evolved from the E-learning, which integrate and compromise traditional collective teaching and the individual learning with ICT. Some scholars have researched and distincted the connotations of hybrid learning from learning theory cluster. According to the actual practice of us, using Joomla to carry out hybrid learning has four characteristics: flexible teaching organization forms and learning ways, flexible teacher-student relationship, flexible application of educational resources, flexible evaluation methods. From the perspective of theory and practics, this paper discusses the building and application of hybrid learning platform based on Joomla, which has a certain degree of practical value and popularizing significance.

References

1. Yi, W.: Design, Development and Application of Blended Learning System. Open Education research (8) (2005)
2. Gabriel, M.: A Modern Tool for Viewing the Learning Resources. In: Hybrid Learning and Education (2009)
3. Liotsios, K., et al.: Blended Learning Technologies in Lifelong Education: Lessons Learned from a Case Study. In: Innovative Approaches for Learning and Knowledge Sharing (2006)

Improving the Quality of Search Results by Eliminating Web Outliers Using Chi-Square

G. Poonkuzhali, R. Kishore Kumar, and R. Kripa Keshav

Department of Computer Science and Engineering,
Rajalakshmi Engineering College, Anna University-Chennai,
Tamil Nadu, 602 105, India
poonkuzhali.s@rajalakshmi.edu.in,
{rskishorekumar,kripa_keshav}@yahoo.co.in,

Abstract. Today, the World Wide Web contains several billions of information and is still growing at a very faster rate as most of the people use the internet for retrieving interesting document. But most of the time, they lose their temper by getting lot of insignificant document even after navigating several links. Thus developing user friendly tool for retrieving the relevant content without accessing the complete data on the outset has become an important concern among the Web mining research communities. In this paper, a mathematical approach is developed for mining this outlaid content which focuses on eliminating irrelevant information. This novel approach uses chi square test for retrieving relevant information. Elimination of irrelevant information during a searching process improves the quality of search engines further. The results show that the system easily provides relevancies and delivers dominant text extraction, supporting users in their query to efficiently examine and make the most of available web data sources.

Keywords: chi-square test, degrees of confidence, relevant, web document, web outliers.

1 Introduction

The past few years have observed the drastic development of the World Wide Web (WWW). Information is being increasingly accessible on the web. The performance and scalability of the web engines face considerable problems due to the presence of enormous amount of web data. The expansion of internet has resulted in problems for the Search engine owing to the fact that the flooded search results are of less relevance to the users. Any one of the subsequent features: different character sets, formats, and inclusions of advertisement or current date may be the reason behind the dissimilarity among identical pages served from the same server. Web crawling is employed by the search engines to populate a local indexed repository of web pages which is in turn utilized to answer user search queries. Business has become more

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 557–565, 2011.

proficient and fruitful owing to the ability to access contents of interest amidst huge heaps of data.

Web mining consists of Web usage mining, Web structure mining, and Web content mining. Web usage mining refers to the discovery of user access patterns from Web usage logs. Web structure mining tries to discover useful knowledge from the structure of hyperlinks. Web content mining aims to extract/mine useful information or knowledge from web page contents. Web Content Mining applies the concepts of data mining and knowledge discovery to retrieve more specific data [1][14]-[16]. Some of the areas of doing research in web content mining is listed below:

- *Structured Data Extraction*
- *Unstructured Text Extraction*
- *Web Information Integration and Schema matching*
- *Building Concept Hierarchies*
- *Segmentation and Noise Detection*
- *Opinion extraction*

This paper focuses on segmentation and detection of noise issue, which implies outliers mining. Generally, Outliers are observations that deviate so much from other observations to arouse suspicions that they might have been generated using a different mechanism or data objects that are inconsistent with the rest of the data objects. Most of the web content mining algorithms works on finding frequent patterns while neglecting less frequent ones namely outliers. Web Content Outliers are web document that show significantly different characteristics than other web documents taken from the same category. Outliers identified in web data are referred to as *web outlier.* This paper focus on Web Content Outliers Mining for extracting insignificant web content through the mathematical approach based on chi-square test.

The proposed system uses chi-square test algorithm to find the web document which are assumed to be more explicit to the user query. Next comparison between the calculated value and the degrees of confidence value at 95% level of significance is done. If the calculated chi-square value is less than tabulated chi-square distribution value, then that document is considered to be significant with the other document which in turn implies it is relevant to the user query. If the chi-square value is greater than chi-square distribution (table value), then that document is considered to be insignificant with the other document which in turn implies it is irrelevant to the user query. Elimination of insignificant web document results in retrieval of interesting web documents to the user.

Outline of the paper:

Section 2 gives the overview of the related works Section 3 presents the overview of the Architectural design of the proposed system. Section 4 presents the algorithm for retrieving relevant web documents using test hypothesis. Section 5 gives the Experimental results. Section 6 presents conclusion.

2 Related Works

Ali et al[1] presents an overview of the major developments in the area of detection of Outliers in numerical datasets. Anguilli et al[2] proposes, a new definition of distance-based outlier and an algorithm, called HilOut, designed to efficiently detect the top n outliers of a large and high-dimensional data set. Breunig et al [4] introduced a new method for finding outliers in a multidimensional dataset through density based approach which uses a local outlier (LOF) for each object in the dataset, indicating its degree of outlier-ness..Bing et al [3] presents characteristics of web and various issues on web content mining. This special issue of SIGKDD Explorations brings together some of the latest research results in this new and exciting field, which continues to make significant impact on real-world applications. Malik Agyemang et al establish the presence of outliers on the web and designed a framework for mining web content outliers using full word matching assuming the existence of domain dictionary. The above authors developed the work with n-gram techniques for partial matching of strings with domain dictionary[10]-[12]. Malik Agyemang et al. enhanced the same work without domain dictionary. Based on the above ideas, Malik Agyemang et..al prolonged the work by presenting HyCOQ which a hybrid algorithm that draws from the power of n-gram based and word based system[13]. There is a remarkable improvement in recall with hybrid documents compared to using raw words and n-grams without a domain dictionary still it covers mining only structured web documents. G.Poonkuzhali et al presented the mathematical approach based on set theory for mining web content outliers [5]. The above authors extended their work based on signed approach using organized domain dictionary for retrieving relevant web document from both structured and unstructured documents.[6]-[7]. G.Poonkuzhali et al proposed a Statistical approach based on proportions using t-test for retrieving the relevant documents [8]. K.Thiagarajan et al. implemented weighted graph approach of trust reputation management through signed concept which can also be applied for retrieving the relevant content [9].

3 Design of the Proposed System

In this algorithm, web documents are extracted based on the user query. The extracted documents are pre-processed for making the further process less complicated. The pre-processing step involves: removal of stop words, stemming and tokenization. Followed by this, term frequency for the words present in the domain dictionary is computed for the i^{th} and j^{th}($j = (i+1)^{th}$) documents. Then, related words from the above documents along with their term frequencies are retrieved for performing chi squaring testing (χ^2). Finally, χ c value is compared with the tabulated Chi-Square value at 95% of level of significance corresponding to degrees of confidence. If the calculated value is lesser or equal to 3.841 then both that documents are considered as relevant documents otherwise they are implied as irrelevant documents. The above process is repeated for all the remaining documents.

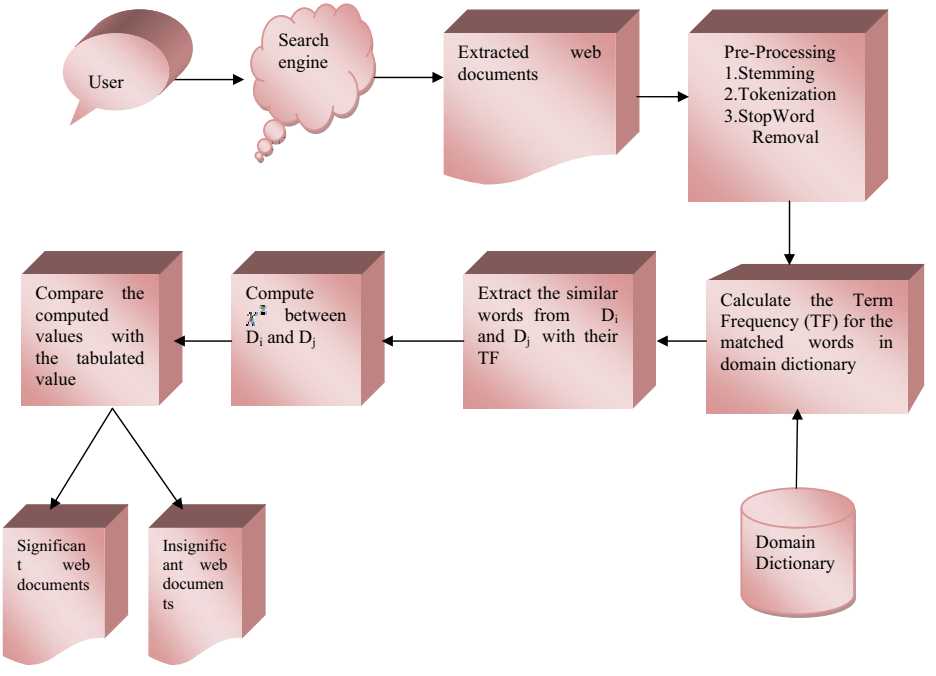


Fig. 1. Architectural Design

4 Algorithm for Retrieving Relevant Document through Test Hypothesis

Input : Web document.
Method: Chi-Square Method.
Output: Extraction of relevant web document.

Step1: Extract the input web document D_i where $1 \leq i \leq N$. Let N be the total number of input documents extracted.

Step 2: Pre-process the entire extracted document.

Step 3: Initialize i=1.

Step 4: Initialize j=i+1.

Step 5: Consider the document D_i and D_j.

Step 6: Extract the words matched with the domain dictionary for further processing.

Step 7: Let N_1 be the total number of words in D_i and N_2 the total number of words in D_j.

Step 8: Perform the Chi Square test through thefollowing steps:

i) Extract the similar words exists between D_i and D_j along with their term frequencies. Let A be the similar words in D_i and B be the similar words in D_j.

ii) Calculate B and D where B= N_1-A and D = N_2 - D.
Now the Observed Values (O) namely A, B, C and D are obtained.

iii) Compute M_1, M_2 and N:
M_1=A+C, M_2=B+D, N=N_1+N_2.

iv) Calculate P, Q, R and S:

$$P = \frac{(M1 * N1)}{N},$$

$$Q = \frac{(M2 * N1)}{N},$$

$$R = \frac{(M1 * N2)}{N},$$

$$S = \frac{(M2 * N2)}{N}.$$

Now the Expected values (E) namely P, Q, R and S are obtained corresponding to A, B, C, and D.

v) Finally, perform Chi-Square(χ^2) Test statistics:

$$\chi^2 = \frac{\sum ((O - E)^2)}{E}$$

Step 9: Calculate the degree of confidence (ndf) as (n-1)*(m-1) where n represents total number of rows and m represents number of columns.

Step10: Get Chi Square distribution value at 95% level of significance corresponding to the degrees of Confidence value .

Step 11: If the value of Chi-Square is less than Chi-Square distribution then
The Documents D_i and D_j are Relevant.
Else
The Documents D_i and D_j are Irrelevant.

Step 12: Increment j, and repeat from step 5 to step 11 until $j \leq N$.

Step 13: Increment i, and repeat from step 4 to step 12 until $i < N$.

5 Experimental Results

Here 5 web documents listed in table1 are taken for test study. Initially these documents are pre-processed and then the term frequencies for the similar words taken for the first two documents are computed. Followed by that, the statistical test hypothesis using chi-square test is applied for those two documents to check the relevancy between them. Similarly, the relevancy for the remaining documents is computed. In this approach the degrees of Confidence at 95% level which holes the value 3.841 is obtained from the Statistical table. The Statistical test value for the input documents is computed in table 2 and relevancy among all the documents is represented in fig. 2.

Table 1. Input documents

D.No	Document Name
D1	Wcm.pdf
D2	Page Content rank an approach to the web content mining.pdf
D3	Neural Analysis.pdf
D4	Deep_WCM.pdf
D5	Medical Mining.pdf

Table 2. Experimental results

	D1	**D2**	**D3**	**D4**	**D5**
D1	*	1.7161	5.4290	0.71217	9.0079
D2	*	*	22.0696	3.2483	23.2509
D3	*	*	*	8.3902	4.3214
D4	*	*	*	*	11.9084
D5	*	*	*	*	*

The above table shows the calculated values of the document after applying the algorithm. Using these values, a graph has been plotted for D_1 against remaining documents. Similarly, the process is repeated for the rest of the documents.

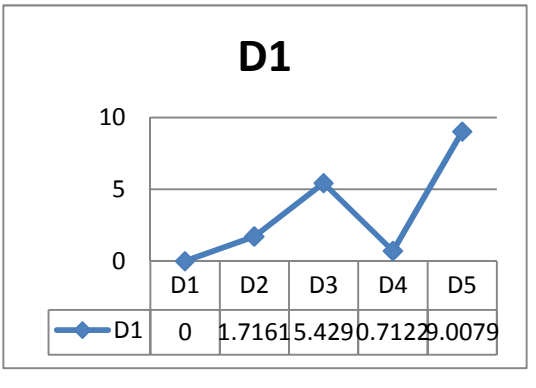


Fig. 2. Graphical representational of experimental results

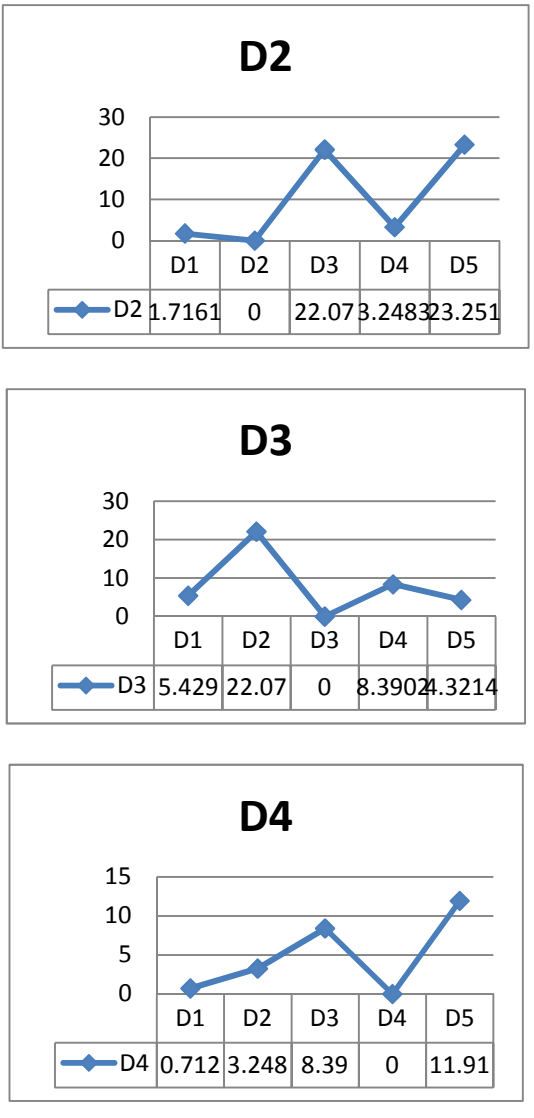


Fig. 2. (*continued*)

From the above figures, it is clear that D2 and D4 are relevant to D1 as the values lie below 3.841.

6 Conclusion and Future Works

Web mining is a growing research area in the mining community. Retrieving relevant content from the web is a very common task. However, the results produced by, most

of the search engines do not essentially produce result that is best possible providing to the user needs. This paper proposes Chi-Square test with 95% degrees of confidence for retrieving relevant web documents from structured as well as unstructured documents. The quality of search results obtained through this approach is very extraordinary and perfect. Future work aims at improvising the chi-square test to increase the precision. Also comparative study with other mathematical models has to be done.

Acknowledgment

The authors would like to thank Prof. Ponnammal Natarajan worked as Former Director – Research, Anna University- Chennai, India and currently an Advisor, (Research and Development), Rajalakshmi Engineering College and Dr. K.. Ravi, Associate Professor, Department of Mathematics, Sacred Heart College- Tirupattur, India for their intuitive ideas and fruitful discussions with respect to the paper's contribution.

References

1. Ali, S., Hadi, A.H.M., Imon, R., Werner, M.: Detection of outliers Overview. Wiley Interdisciplinary Reviews: Computational Statistics 1(1), 57–70 (2009)
2. Angiulli, F., Pizzuti, C.: Fast Outlier Detection in High Dimensional Spaces. In: Elomaa, T., Mannila, H., Toivonen, H. (eds.) PKDD 2002. LNCS (LNAI), vol. 2431, pp. 15–27. Springer, Heidelberg (2002)
3. Liu, B., Chang, K.C-.C.: Editorial: Special issue on Web Content Mining. SIGKDD Explorations 6(2)
4. Breunig V., M.M., Kriegel, H.-P., Ng, R.T., Sander, J.: LOF: Identifying Outliers in Large Dataset. In: Proc. of ACM SIGMOD 2000, Dallas, TX (2000)
5. Poonkuzhali, G., Thiagarajan, K., Sarukesi, K.: Set theoretical Approach for mining web content through outliers detection. International Journal on Research and Industrial Applications 2, 131–138 (2009)
6. Poonkuzhali, G., Thiagarajan, K., Sarukesi, K., Uma, G.V.: Signed approach for mining web content outliers. Proceedings of World Academy of Science, Engineering and Technology 56, 820–824
7. Poonkuzhali, G., Sarukesi, K., Uma, G.V.: Web Content Outlier Mining Through Mathematical Approach And Trust Rating. In: The Book " Recent Researches In Applied Computer And Applied Computational Science", Included in ISI/SCI Web of Science and Web of Knowledge,Venice, Italy, pp. 77–82 (2011)
8. Poonkuzhali, G., Kishore Kumar, R., Kripa Keshav, R., Sarukesi, K.: Paper titled "Statistical Approach for Improving the Quality of Search Engine". In: The Book " Recent Researches In Applied Computer And Applied Computational Science", Included in ISI/SCI Web of Science and Web of Knowledge,Venice, Italy, pp. 89–93 (2011)
9. Thiagarajan, K., Raghunathan, A., Natarajan, P., Poonkuzhali, G., Ranjan, P.: Weighted Graph Approach for Trust Reputation Managements. In: International Conference on Intelligent Systems and Technologies, vol. 56, pp. 830–836. Published in Proc. Of World Academy of Science and Technology (2009)

10. Agyemang, M., Barker, K., Alhajj, R.S.: Framework for Mining Web Content Outliersb. In: ACM Symposium on Applied Computing, Nicosia, Cyprus, pp. 590–594 (2004)
11. Agyemang, M., Barker, K., Alhajj, R.S.: Mining Web Content Outliers using Structure Oriented Weighting Techniques and N-Grams. In: ACM Symposium on Applied Computing, Santa Fe, New Mexico, pp. 482–487 (2005)
12. Agyemang, M., Barker, K., Alhajj, R.S.: WCOND –Mine: Algorithm for detecting Web Content Outliers from Web Documents. In: IEEE Symposium on Computers and Communication (2005)
13. Agyemang, M., Barker, K., Alhajj, R.S.: Hybrid approach to web content outlier mining without query vector. In: Tjoa, A.M., Trujillo, J. (eds.) DaWaK 2005. LNCS, vol. 3589, pp. 285–294. Springer, Heidelberg (2005)
14. Kosala, R., Blockeel, H.: Web Mining Research: A Survey. ACM SIGKDD (July 2000)
15. Campos, R., Dias, G., Nunes, C.: WISE: Hierarchical Soft Clustering of Web Page Search Results based on Web Content Mining Techniques. In: International Conference on Web Intelligence, IEEE/WIC/ACM (2006)
16. Pol, K., Patil, N., Patankar, S., Das, C.: A Survey on Web Content Mining and Extraction of Structured and Semistructured data. In: First International Conference on Emerging Trends in Engineering and Technology (2008)

Segmentation of Fiber Image Based on GVF Snake Model with Clustering Method

Xin Li[1], Li Yao[2], and Yan Wan[2]

[1] The Third Research Institute of Ministry of Public Security Shanghai, China
[2] School of Computer Science and Technology Donghua University Shanghai, China
nclixin@sina.com, {yaoli,winniewan}@dhu.edu.cn

Abstract. In the fiber image analysis system, correctly segmenting fiber from fiber micrograph is critical for fiber feature extraction and further identification. In this paper, the GVF snake model with the initial contour obtained by contour tracking method based on K-means clustering segmentation is proposed for fiber segmentation. Firstly, the K-means clustering method is used to obtain the initial coarse contour of fiber, and then the GVF Snake algorithm is applied to calculate the accurate fiber contour. Due to the noise of fiber image, some fiber contours have burrs, which can be removed by contour tracking method. Experiment result shows that this algorithm can obtain the boundaries of desired object from fiber image effectively and accurately, meanwhile, the new method expands apply area of the snake model to process the complicated image.

Keywords: GVF snake model; fiber image; K-means clustering; contour tracking.

1 Introduction

Fiber mixing ratio plays an important role in determining the ultimate quality on the final product, and thus needs to be monitored quantitatively for quality inspection. Traditionally, fiber detecting depends on manual or semi-manual which has time-consuming, low precision, data poor stability defects. Currently, microscopic image analysis is one of the most common ways for fiber content measurements. It can greatly reduce the workload of inspectors, relieve the recognition difficulty and improve work efficiency when computer image processing technology used in automatic fiber identification system 1. Automated identification of fiber is critically important for fiber quality inspection and it has six steps which are sample preparation, image capturing, image pre-processing, image segmentation, feature extraction and fiber identification. After sample preparation, fiber images can be got from image capturing, and nonuniform illumination is removed by image pre-processing. Single fiber is segment from image background by fiber segmentation. According the fiber's feature extracted by feature extraction, fibers are finally classified and fiber mixing ratio is calculated.

Fiber segmentation is a task critical to automatic fiber identification, because the result of fiber identification is based on the accuracy of fiber boundaries detected by fiber segmentation algorithm. Many researchers have made great efforts to present image segmentation algorithms, which can be categorized into two types: edge based and threshold based. Edge-based segmentation which puts emphasis on detecting

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 566–574, 2011.

significant gray-level changes near region boundaries, always uses differential operator like Sobel operator, Log operator and Canny operator to detecting image contour, then gets the profile of object with the algorithm of tracking or extracting image contour. Threshold-based segmentation is widely applied by its advantage of simplicity and effectiveness, but the proper threshold value has to be carefully chosen to achieve the best results. Due to the impact of little gray-scale differences between background and foreground of fiber image, it is difficult to select a proper threshold value. Fig.1 compares the results of the edge-based segmentation (Sobel operator (b), Log operator (c) and Canny operator (d)) and threshold-based segmentation (Ostu algorithm (e), one-dimensional maximum entropy method (f) and minimum error thresholding method (g)) when they were applied to fiber image. Due to the noise of image and gray-level changes not great enough for detecting boundaries, these methods tended to generate pseudo, dual and/or broken edges, leading to false measurements and identification.

In recent years research on segmenting image with deformable models is hot. Active contours, or snakes algorithm, have been widely studied and applied in image analysis. Snake was originally introduced in 1987 by Kass 2, and has a number of attractive properties that make it useful in edge detection, shape modeling, segmentation, and motion tracking. Snakes are curves defined within the image domain that can move under the influence of internal forces coming from within the curve itself and external forces computed from the image data. The internal and external forces are defined so that the snake will conform to the object's boundary. Compare with the above-mentioned segmentation algorithms, the greatest advantage of snake model is that it can calculate a consecutive and single-pixel edge.

However, the traditional snake model needs goods selections of the initial contour which often been done by human, and the concave boundary cannot be detected accurately. Various improvements have been proposed. Eviatar et al. 3 have presented an algorithm to solve the problem about the boundary concavities, but the initialization of original snake still need manual work. Cohen 4 has imported external force to expand and shrink the active contour, but it is not easy to ensure the force. Yuen 5 proposed an automatic initialization snake algorithm for multi-object segmentation. Nevertheless, their algorithm is limited to some specifically distributed objects, for example those spreading around the center of gravity. Xu and Prince 6-8 have proposed a new external force model for active contours and deformable surfaces, which called the Gradient Vector Flow (GVF) field. The field is calculated as a diffusion of the gradient vectors of a gray-level or binary edge map derived from the image. GVF snake has a large capture range and is able to move into concave boundary regions compared with the traditional snake. However, The GVF snake is apt to be misdirected by the noise, and hardly detects the correct contours when the initial contour is far away from the true boundary.

For detecting suitable initial contour, clustering algorithm is a good choice. K-means clustering algorithm is a simple clustering method with lower computational complexity. However, the K-mean clustering algorithm has several weaknesses. It is sensitive to initial clustering center, and different initial centers often correspond to different clustering results. Besides, K-means algorithm for the noise and isolated points in the data sets is very sensitive. In this paper, the K-means method is improved to detect the fiber's area automatically. This area should be accurate enough so it can be the initial contour of GVF Snake algorithm.

2 GVF Snake Model

A. Traditional Snake Model

The Snake presented by Kass is a model of a deformable curve or contour (if closed) composed of abstract elastic materials. Active contours, or snakes, are curves defined within an image domain that can move under the influence of internal forces within the curve itself and external forces derived from image data. Energy in this active contour model is represented by two force terms: E_{int} and E_{ext}. A traditional snake is a curve $\boldsymbol{x}(s) = [x(s), y(s)], s \in [0,1]$, which moves through the spatial domain of an image to minimize the energy function. The energy functional to be minimized is defined as 8:

$$E_{energy} = \int_0^1 \frac{1}{2}\left(E_{int}(\boldsymbol{x}(s)) + E_{ext}(\boldsymbol{x}(s))\right) ds \tag{1}$$

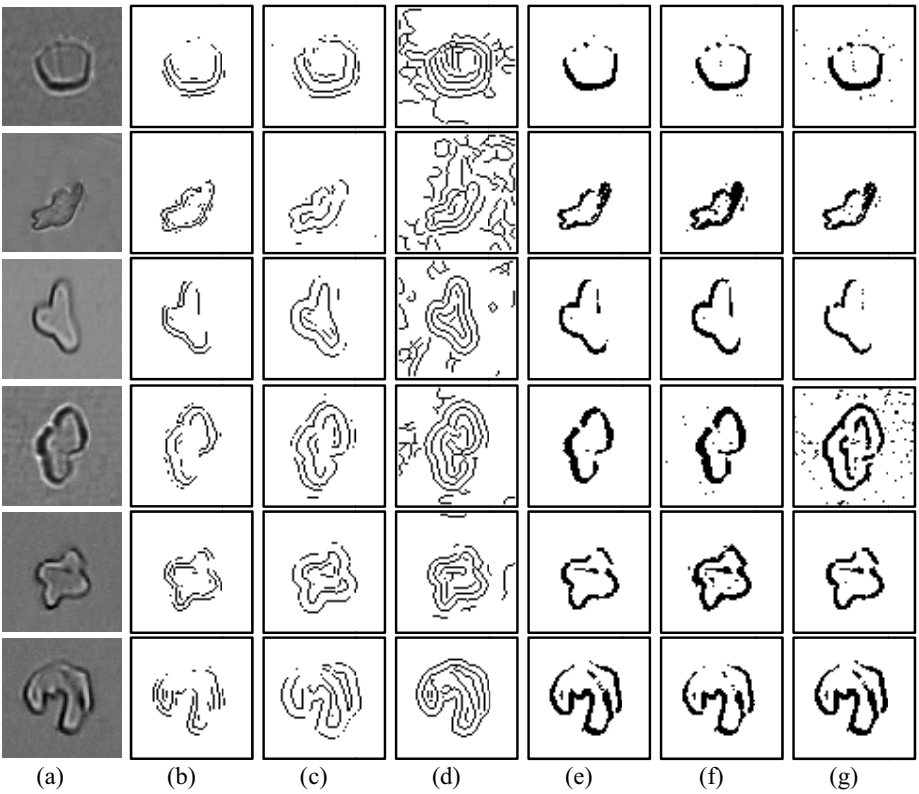


Fig. 1. (a) Original fiber images, the comparison among the segmentation results by (b) Sobel operator, (c) Log operator, (d) Canny operator, (e) Ostu algorithm, (f) one-dimensional maximum entropy method, and (g) minimum error thresholding method

E_{int} represents the internal potential energy of the snake, which is a function of both bending and stretching forces applied to the snake. E_{ext} gives rise to external constraint forces. The internal potential energy of the snake E_{int} is defined as follows:

$$E_{int}(\boldsymbol{x}(s)) = [\alpha|\boldsymbol{x}'(s)|^2 + \beta|\boldsymbol{x}''(s)|^2] \tag{2}$$

where α is defined as the flexible parameter while the β is defined as the rigid parameter that can control the stretch and the bend of a snake. $\boldsymbol{x}'(s)$and $\boldsymbol{x}''(s)$ denote the first and second derivatives of $\boldsymbol{x}(s)$ with respect to s.

E_{ext} is derived from the image so that it takes on its smaller values at the features of interest, such as boundaries, and here is defined as:

$$E_{ext}(x,y) = -|\nabla(G_\sigma(x,y) * I(x,y))|^2 \tag{3}$$

where $I(x,y)$is the image of gray level which viewed as a function of continuous position variables (x,y). $G_\sigma(x,y)$ is a two-dimensional Gaussian function with standard deviation, σ and ∇ is the gradient operator.

A final solution will be obtained by minimizing total snake energy which must satisfy the Euler equation:

$$\alpha\boldsymbol{x}''(s) - \beta\boldsymbol{x}(s) - \nabla E_{ext} = 0 \tag{4}$$

To find a solution to Eq. (4), the snake is made dynamic by treating $\boldsymbol{x}$ as function of time t as well as s, i.e., $\boldsymbol{x}(s,t)$. Then, the partial derivative of $\boldsymbol{x}$ with respect to t is then set equal to the left-hand side of Eq. (4) as follows:

$$x_t(s,t) = \alpha x''(s,t) - \beta x''''(s,t) - \nabla E_{ext} \tag{5}$$

The term $\boldsymbol{x}_t(s,t)$ dissolves when the solution $\boldsymbol{x}(s,t)$ stabilized, and the solution of Eq. (4) can be achieved. The dynamic Eq. (5) can be viewed as a gradient descent method designed to solve Eq. (1). That is to say we can obtain the object contours.

The snake is deformed due to external forces that attract it towards salient features of the image, and internal forces which try to preserve the smoothness of the shape of the contour. Although traditional snake has found many applications, it is essentially weak in two aspects. Firstly, in general, the initial contour must be fairly close to the true boundary and construction of initial contour often requires human interaction. The second problem is that the active contour hard to progress into concave boundary positions. Consequently, a new snake model, Gradient Vector Flow snake, has been proposed by Xu and Prince 6 to achieve better object segmentation and remedied both of the shortcomings of the traditional snake.

B. Gradient Vector Flow Snake

The basic idea of the GVF snake is to extend influence range of image force to a large area by generating a GVF field. The GVF field is computed from the image.

The GVF field has been defined as a new static external force field $\boldsymbol{v}(x,y) = [u(x,y), v(x,y)]$, and is used to replace the potential force $-\nabla E_{ext}$ in (5) with $\boldsymbol{v}(x,y)$, yielding 7:

$$\boldsymbol{x}_t(s,t) = \alpha\boldsymbol{x}''(s,t) - \beta\boldsymbol{x}''''(s,t) + \boldsymbol{v} \tag{6}$$

$\boldsymbol{v}$ minimizes the energy function:

$$\varepsilon = \iint \mu\left(u_x^2 + u_y^2 + v_x^2 + v_y^2\right) + |\nabla f|^2 |\boldsymbol{v} - \nabla f|^2 dxdy \tag{7}$$

where μ is a regularization parameter governing the tradeoff between the first term and the second term in the integrand. This parameter should be set according to the amount of noise present in the image (more noise, increase μ). f is an edge map which is derived from the original image, having the property that it is larger near image edges. Resulting from this minimization, when $|\nabla f|$ is large, the second term dominates the integrand, and produces the effect of keeping $\boldsymbol{v}$ nearly equal to∇f. On the contrary, when $|\nabla f|$ is small, the energy is dominated by sum of the squares of the partial derivatives of the vector field.

Using the calculus of variations, we can get the GVF by solving the following Euler-Lagrange equations:

$$\mu\nabla^2 u - (u - f_x)\left(f_x^2 + f_y^2\right) = 0 \tag{8a}$$

$$\mu\nabla^2 v - \left(v - f_y\right)\left(f_x^2 + f_y^2\right) = 0 \tag{8b}$$

where ∇^2 is the Laplacian operator.

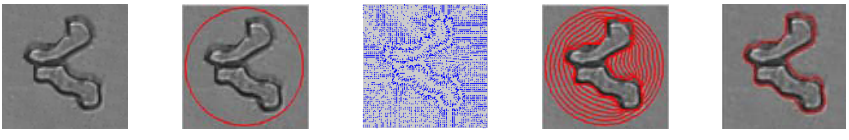
To find a solution to Eq. (7), $\boldsymbol{u}$ and $\boldsymbol{v}$ are treated as functions of time t and solving

$$u_t(x,y,t) = \mu\nabla^2 u(x,y,t) - \left(u(x,y,t) - f_x(x,y)\right) \cdot\left(f_x(x,y)^2 + f_y(x,y)^2\right) \tag{9a}$$

$$v_t(x,y,t) = \mu\nabla^2 v(x,y,t) - \left(v(x,y,t) - f_y(x,y)\right) \cdot\left(f_x(x,y)^2 + f_y(x,y)^2\right) \tag{9b}$$

After computing the GVF field, take it as external force and apply it into the iterative Eq. (6). Using iterative calculation, we can find the solutions when the time variable is reaching a steady state.

Fig. 2 shows an example of the segmentation process of fiber image with GVF snake. Fig. 2(b) shows the initial border of Fig. 2(a) using GVF snake model, Fig. 2(c) shows the GVF snake force field, then Fig. 2(d) gives the processing of obtaining the fiber contour with GVF snake algorithm and the final result in Fig. 2(e).



(a) (b) (c) (d) (e)

Fig. 2. (a) Original fiber image; (b) initial border; (c) GVF force field; (d) the processing of obtaining the fiber contour with GVF snake model; (e) convergence using a GVF snake

In general, advantages of GVF snake over a traditional snake are its insensitivity to initialization and ability to move into concave boundary regions. Furthermore, the

GVF field greatly increases the capture range of the snake. If the initial contour is far away from the true boundary, it would increase the iterations of contour at the same time much more computation is needed. To deal with these problems, we have developed an improved GVF snake. This new method is presented in detail in the following section.

3 Improve GVF Snake Model

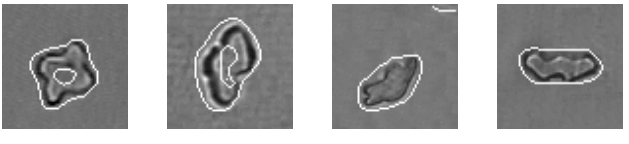
C. K-Means Clustering Algorithm

Clustering is the process of partitioning or grouping a given set of patterns into disjoint clusters. It has been a widely studied problem in a variety of application domains including face recognition, fingerprint recognition and biomedical image processing. Researchers have implemented numerous clustering algorithms. One of the most important and widely used algorithms is the K-means clustering. It was first introduced by McQueen 9 for data classification and analysis. The K-means method is numerical, unsupervised, non-deterministic and iterative and commonly used in computer vision as a form of image segmentation.

K-means clustering is an algorithm of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. Given a set of observations $S=(x_1, x_2, \ldots, x_n)$, where each observation is a d-dimensional real vector, the K-means clustering aims to partition the n observations into k ($k < n$) sets $S=(S_1, S_2, \ldots, S_k)$ so as the metric W in Eq. (10) reaches minimum.

$$W = \sum_{i=1}^{k} \sum_{x_j \in S_i} \left| x_j, w_i \right|^2 \tag{10}$$

where w_i represents the center points or mean point of all the data x_j in S_i.



(a) (b) (c) (d)

Fig. 3. White lines depicted the results of K-means clustering on our fiber images

In the image processing, the data set would be a set of pixel vectors. Hence, each pixel of the image will be classified into a cluster. The K-means clustering algorithm selects k points as initial cluster centers, and then the iterative operation begins. Different selection of initial point can achieve different clustering result. For the reduction of the clustering result's dependence on the initial value and the improvement of the initial cluster centers can be achieved by the search algorithm of the cluster center. In our paper, if the k value is too large, it would result in fiber image

over-segmented. Otherwise, it would cause the results of image segmentation too rough. They both affect the following image process. During our algorithm, we want to obtain a coarse contour of the fiber with K-means clustering which can describe the fiber's general profile feature. According to the characteristics of fiber image, we select $k \in [4, 8]$ in our paper.

Fig. 3 shows the edge detection results of various fiber images by K-means clustering method. It shows that not only the noise is suppressed, but also the true image edges can be found, though the edge detected by this method is coarse.

Even though K-means clustering algorithm cannot obtain the precise contour of fiber, it can achieve a general profile of fiber. In our algorithm, we use contour tracking method to get the longest edge of the image based on the result of K-means clustering for the initial contour of GVF snake. With the initial contour, GVF snake can capture the accurate fiber contour.

D. Contour Tracking Algorithm

The initial contour of GVF snake model could directly influence the GVF field and the time of convergence. It is essential for us to setting the initial contour with a simple and fast method. K-means clustering algorithm is an uncomplicated method with lower computational complexity and it can obtain a general profile which close to the real contour. If we choose a proper edge of K-means clustering's results as the initial contour of GVF snake model, it would reduce the iterations of GVF snake method and easily get the desired fiber contour, so then enhance the efficiency of segmentation.

We obtain the edge image by the K-means clustering algorithm. Then we use the contour tracking method to get the longest edge for initial contour of GVF snake model. With this initial contour, GVF snake can capture the object boundaries. Fig. 4 shows the processing of fiber image segmentation with GVF snake model based on K-mean clustering and tracking contour algorithm. Fig. 4(b) is segmentation result of fiber image in Fig. 4(a) by K-means clustering. Then in Fig. 4(c) we obtain the longest contour with tracking contour method for GVF snake model's initial contour. Fig. 4(d) shows the processing of obtaining the contour with GVF snake model, and Fig. 4(e) is the final result.

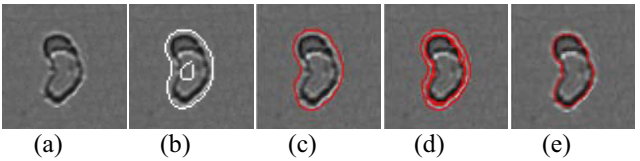
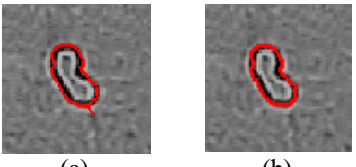


Fig. 4. (a) Original fiber image (b) segmentation result of K-means clustering (c) the initial contour of GVF snake model (d) the processing of obtaining the fiber contour with GVF snake model (e) final fiber contour by our proposed algorithm

However, there are a lot of noises when fiber image is taken, which will lead to burrs, if we directly get the initial contour of GVF snake model with tracking contour method, there would be burrs in some fiber images. The result with burrs could influence extraction of contour characteristic such as fiber's area, fiber's perimeter and so on. In our paper, we remove burrs by tracking the contour. During the tracking

process, if it arrives at the endpoint of the burr and there is no curve pixel without being tracked in its 8-connected neighborhood, we have to return and track the burr pixels again. Thus, tracking time of the burr pixels is at least once more than that of curve pixels. Even if the curve is tracked several times, we will track the crossing points of the curve and the burr at least once more than curve pixels. Generally speaking, the minimum tracking times is the tracking times of curve pixels. However, if we directly delete the pixels with different tracking times from that of the curve pixels, the starting point and the crossing points will be mistaken for the burr pixels and then be deleted, which will affect the connectivity of the closed curve. Accordingly, starting point is processed particularly. Stating point should be retained when it isn't on burr, and its tracking times would be subtracted one. Removing the starting point is same to other pixels. For a crossing point, if there is a pixel in its neighborhood, the pixel cannot be deleted if there is no curve pixel in its front or back position. Fig. 5(a) shows a fiber contour has burrs and the Fig. 5(b) is the result of removing burrs of Fig. 5(a) by tracking method.



(a) (b)

Fig. 5. (a) The fiber contour with burrs (b) after removing burrs

4 Experiments and Conclusion

Fig. 6 shows the results of segmentation of fiber images with K-means clustering and GVF snake model. We can see our method easily converges to the true boundary and solves the drawbacks of traditional segmentation algorithm, and then a single and consecutive curve can be obtained. This curve keeps the accurate geometrical feature of fiber edge, which is critical for fiber identification.

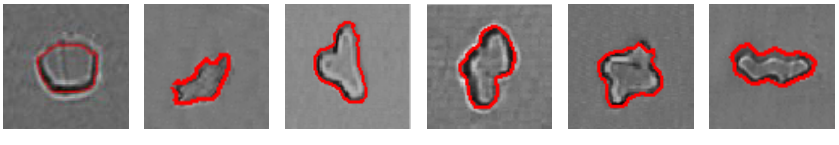


Fig. 6. The edge result detected by our proposed algorithm

In this paper, we have present a new fiber image segmentation algorithm combining GVF snake, K-means clustering and tracking contour algorithm. At first, the initial coarse contour of fiber is calculated by K-means clustering. Then the tracking contour algorithm is used to obtain the longest edge for GVF snake model's initial contour. With the initial contour, GVF snake can calculate the accurate fiber boundaries. Our method solves the problem that the edge based on k-means clustering is not precise and it can cut down the iterations of GVF field and the number of contour approaching.

Therefore, the segmentation of fiber image with K-means clustering and GVF snake model decreases the computation greatly and makes the snake more robust to noise. Experiments indicate that the new algorithm expand the apply area of snake model to process the complicated images.

References

1. Wan, Y., Yao, L., Xu, B., Zeng, P.: Shaped Fiber Identification Using A Distance-Based Skeletonization Algorithm. Textile Research Journal 80(10), 958–968 (2010)
2. Kass, M., Witkin, A., Terzopouls, D.: Snake: active contour models. International Journal of Computer Vision 1(4), 321–331 (1987)
3. Davison, N.E., Eviatar, H., Somorjai, R.L.: Snakes simplified. Pattern Recognition 33(10), 1651–1664 (2000)
4. Cohen, L.D., Cohen, I.: Finite-element methods for active contour models and balloons for 2-D and 3-D images. IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1131–1147 (1993)
5. Yuen, P.C., Feng, G.C., Zhou, J.P.: A contour detection method: Initialization and contour model. Pattern Recognition Letters 20(2), 141–148 (1999)
6. Prince, J.L., Xu, C.: A new external force model for Snakes. In: Proc. IEEE Conference on Computer Visition Pattern Recognition, pp. 30–31 (1996)
7. Xu, C., Prince, J.L.: Snakes, shapes, and gradient vector flow. IEEE Transactions on Image Processing 7(3), 359–369 (1998)
8. Xu, C., Prince, J.L.: Generalized gradient vector flow external forces for active contours. Signal Processing 71(2), 131–139 (1998)
9. McQueen, J.B.: Some Methods for Classification and Analysis of Multivariate Observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistic and Probability, vol. 1, pp. 281–297. University of California Press, Berkeley (1967)

Adopting a Connection Oriented Private Cloud Desktop to Facilitate Blended Learning

Chin-Yi Yang[1], Ching-Tao Chang[1], Li-Ren Chien[1], and Chun-Yao Wang[2]

[1] Department of Information Management, National Yunlin University of Science and Technology, 123 University Road, Section 3, 64002 Yunlin, Taiwan
[2] Department of Computer Scienc and Information Engineering, National Chung Cheng University, 168, University Rd., Min-Hsiung Chia-Yi, Taiwan.
62102 Heidelberg, Taiwan
rebacayang@gmail.com,
{clj,cyw}@ccu.edu.tw

Abstract. A connection oriented private cloud desktop named Dumbogo is developed to support cloud computing on anytime, anywhere and any device (3A). Cloud computing may be the most important paradigm shift after PC revolution and Internet prevailing. So far as to cloud computing in education, it would have significant impact on pedagogy since the inherited 3A characteristics. In this paper we expose to use Dumbogo as a media which facilitates cloud computing to sustain lesson planning, teaching activity and review. In the pilot research the demo site run Hsin Kuo High School in south Taiwan, we propose a new teaching approach under Dumbogo in a blended learning environment with learning style consideration.

Keywords: cloud computing in education, pedagogy, blended learning.

1 Introduction

There is a growing disparity of rich and poor that is evident [1]. Education is the most important mechanism to enhance income and increase an upward flow among social classes, where the education level is strongly related to the resources that students can gain. High quality educational resources are usually hidden behind high tuition payments. Superior educational resources thus become monopolized by higher social classes. On the contrary, the people in lower social classes acquire fewer educational resources, leading to poorer educational achievement. The more they have, the more they get, thus producing an M-shape in educational achievement. Then M-shaped society [2] problems follow soon after. To avoid an M-shaped society, the elimination of M-shaped educational resources becomes a very important issue.

Many frontiers do effort on providing superior educational resources for the masses. Editors of Wikipedia are not motivated by money, but rather by a desire to make a lasting contribution to society. Similarly, the teachers will use the apps provided by the cloud to choose the 'cream of the crop' of educational materials for students. Our resource equally allocation plan starts from campus by private cloud [3].

M. Zhou and H. Tan (Eds.): CSE 2011, Part II, CCIS 202, pp. 575–582, 2011.

Aimed at providing superior educational resources for everybody, we have designed an educational cloud named Dumbogo where students can use cheaper devices to reach the educational resources from anywhere and anytime. We also consider that local schools can't afford expensive mainframes, professional servers and high speed networks with their limited budgets. Our solution is a web-based resource sharing platform. Dumbogo is a virtual desktop [4] that runs on browsers. We use GWT [5] to build user interfaces for apps. Students can use any device that supports a JavaScript browser to reach Dumbogo, including PCs, laptops, notebooks, pads and smart phones. The resources we share consist of document-based teaching materials and online multimedia that can be used in a blended learning classroom. We are currently using Dumbogo at Hsin Kuo High School located in southern Taiwan. We use open-source Ubuntu and Apache Hadoop to abandon costly devices, instead using a local area network of PCs to as data servers for our private cloud. High availability, high scalability and fault-tolerance are achieved by using the cloud's virtualization, which can tolerate such hardware problems.

In the pilot research we propose innovation pedagogy to support lesson planning, teaching activity and review under blended learning environment. Meanwhile, teachers can dispatch suitable learning material to the students considering the learning style.

In subsequent section, literature review work on blended learning environment and learning style will be the following. Then the development of the solution Dumbogo will be set forth. Finally the paradigm shift of pedagogy in blended learning environment will be introduced in the pilot study.

2 Background

Before discussing the practical aspects of Dumbogo, it is important to understand the concepts of blended learning and learning style.

2.1 Blended Learning

Blended learning has been referred to as the ''third generation'' [6] of distance education systems. The first generation was correspondence education which utilized a one-way instructional delivery method, including mail, radio, and television. The second generation was distance education with single technology, such as computer-based or web-based learning. The third generation is blended learning, characterized as maximizing the best advantages of face-to-face learning and multiple technologies to deliver learning.

While finding a great deal of fuzziness in the blended learning literature review, we spotlight on a compromise between the conventional face-to-face sessions and online learning shown as Fig. 1. According to this concept, we developed a Teacher Control Model of a Blended Learning Environment based on a Typical Classroom in Secondary School [7]. The proposed innovation pedagogy could be started from there.

Reasons for using blended methods vary. Osguthorpe and Graham [8] stated that educators adopted blended learning to improve the following areas: pedagogy, access to knowledge, social interaction, personal presence, cost effectiveness and ease of revision. Chung and Davis [9] reported that blended learning extended classroom interactions between the students and the instructor through the use of e-mail

asynchronous discussion boards. They also found that learning technologies provided learners with greater control over the pace of learning, instruction flow, selection of resources and time management. With the capacity to accommodate both self-regulated and interpersonal learners, blended learning was also found effective in addressing diverse learning style [10].

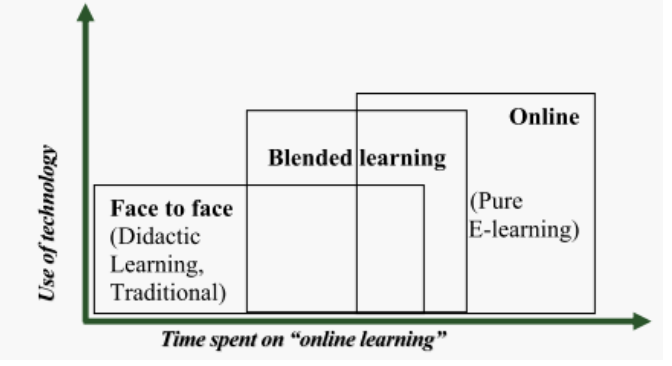


Fig. 1. Conception of Blended Learning [11]

2.2 Learning Style

Researcher in instructional psychology has demonstrated that adapting instructional methods and teaching strategies to accommodate key individual differences including learning style has led to improved performance [12]. The mechanism through which an individual's learning style influences his/her learning is explained by learning style theories. In this section, we will review the definition of learning style and introduce Kolb's learning style inventory (KLSI).

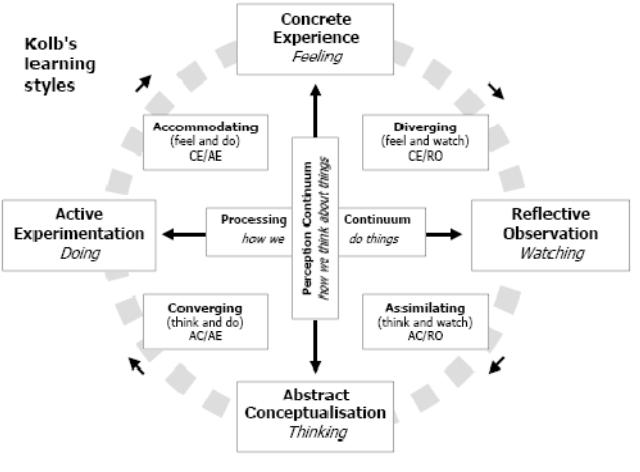


Fig. 2. Learning Style Grid [13]

Kolb's learning style inventory, as defined by [13] and modified in 2005 [14], includes four learning styles:

Converger:
People with this learning style are best at finding practical uses for ideas and theories. They have the ability to solve problems and make decisions based on finding solutions to questions or problems. Individuals with a Converging learning style prefer to deal with technical tasks and problems rather than with social issues and interpersonal issues. These learning skills are important for effectiveness in specialist and technology careers. In formal learning situations, people with this style prefer to experiment with new ideas, simulations, laboratory assignments, and practical applications [14].
Diverger
A person labeled Diverging performs better in situations that call for generation of ideas, such as a brainstorming session. People with a Diverging learning style have board culture interests and like to gather information. They are interested in people, tend to be imaginative and emotional, have broad interests, and tend to specialize in the arts. In formal learning situations, people with the Diverging style prefer to work in groups, listening with an open mind to different points of view and receiving personalized feedback[14].
Assimilator
People with this learning style are best at understanding a wide range of information and putting it into concise, logical form. Individuals with an Assimilating style are less focused on people and more interested in ideas and abstract concepts. Generally, people with this style find it more important that a theory have logical soundness than practical value. The Assimilating learning style is important for effectiveness in information and science careers. In formal learning situations, people with this style prefer readings, lectures, exploring analytical models, and having time to think things through [14].
Accommodator
People with this learning style have the ability to learn from primarily "hands-on experience". They enjoy carrying out plans and involving themselves in new and challenging experiences. Their tendency may be to act on "gut" feelings rather than on logical analysis. In solving problems, individuals with an Accommodating learning style rely more heavily on people for information than on their own technical analysis. This learning style is important for effectiveness in action-oriented career such as marketing or sales. In formal learning situations, people with the Accommodating learning style prefer to work with others to get assignments done, to set goals, to do field work, and to test out different approaches to completing a project [14].

3 Implementation

Aimed at providing superior educational resources for everybody, we have designed an educational cloud where students can use cheaper devices to reach the educational resources from any place and any time. We also consider that local schools can't afford expensive mainframes, professional servers and high speed networks with their limited budgets. Our solution is a web-based resource sharing platform. Dumbogo is a

virtual desktop that runs on browsers. We use GWT to build user interfaces for apps. Students can use any device that supports a JavaScript browser to reach Dumbogo, including PCs, laptops, notebooks, pads and smart phones.

3.1 Story of Dumbogo

We named our system as "Dumbogo" coming from a Disney movie: Dumbo [15] while "Dumbo" is inspired by "Hadoop" [16] which is also a stuffed elephant.

In the beginning, we used Hadoop as distributed file system to manage files through a web-based system. Originally we wanted to provide a web-based environment for users easily accessing HDFS files. Ideas sprang from doing that. When HDFS grew mature, we found we could update the web-based file system into a private connection-oriented cloud virtual desktop. Besides files systems, after authorization, users can have personalized individual and group functions like geolocation, movie playing, text showing and connecting to widgets from the Net. Unlimited possibilities occur when Dumbogo starts flying.

"Go Go Dumbo" is the slogan of our team. We abbreviated "Go Go Dumbo" as "Dumbogo."

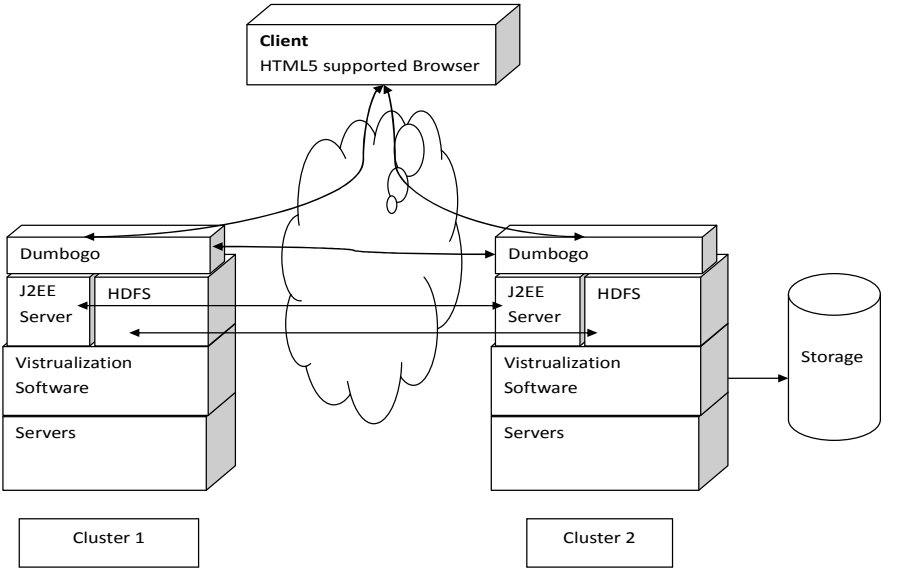


Fig. 3. Framework of Dumbogo

3.2 Framework

Dumbogo is a connection oriented Private Cloud Virtual Desktop OS based on Hadoop Distributed File System (HDFS) with a Transaction Session Management System. A Dumbogo client runs on any device which supports HTML 5 browser, since the Dumbogo server runs as a .war wrapper that can be run on any J2EE Servlet Server. Dumbogo was developed with Google's GWT.

Dumbogo as a Web application was developed under Linux, Web Application Archive (WAR), Hadoop Distributed File System and Google Web Tools (LWHG) to against the traditional LAMP (Linux, Apache, Mssql and PHP).

As figure 3 shows, the Dumbogo is run on clusters. A cluster consists of several severs. The servers are running visualization software. As an example, a cluster could be consisted of six HP DL-380 servers with VMware ESX. Numbers of virtual machines (VMs) were created and managed by the visualization software. The Dumbugo system called the functions of visualization software by RESTful to automatic the management of the visual machine level. For example, Dumbogo will create a new Ubuntu server by VMware function. The new server can be a new data node of HDFS when Dumbogo found the system lacks storage space.

At least one J2EE sever should run on a cluster. The Dumbogo server was packed as Web Application Archive format and can be deployed on any J2EE container (i.e. Tomcat, Glassfish, JBoss … etc.). Dumbogo supports the communication between Dumbogo servers peer to peer. That makes the servers can easily change the information of their users in its cluster. Dumbogo handles Hadoop Distribute File System and saves/retrieves file by Map/Reduce matter.

As figure 4 shows, the Dumbogo client was developed by GWT and run on any HTML 5 supported browser. This makes the users can easy access the resource distributed on HDFS by any devices (i.e. PC, Notebook, Pad and Smart Phone)

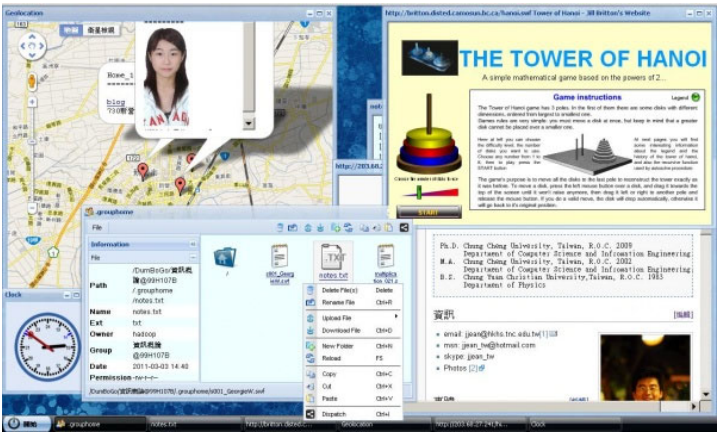


Fig. 4. A Dumbogo client run on HTML5 browser

4 Pedagogy Paradigm Shift

Dumbogo is a virtual desktop the teacher and student can login respectively from remote by browser. The inherited characteristics from cloud computing anyplace, anywhere and anytime can support blended learning environment.

Dumbogo could also be a resource container which composes of document-based teaching materials and online multimedia. There are many kinds of resource under a topic in teaching. Teachers always choose the ‘cream of the crop’ of educational

materials to students. What are the 'cream of the crop' materials should be considered from students. Individual difference should be considered in the teaching process. In this pilot research we will focus on learning style.

In this section we propose innovation pedagogy under Dumbogo to facilitate teaching.

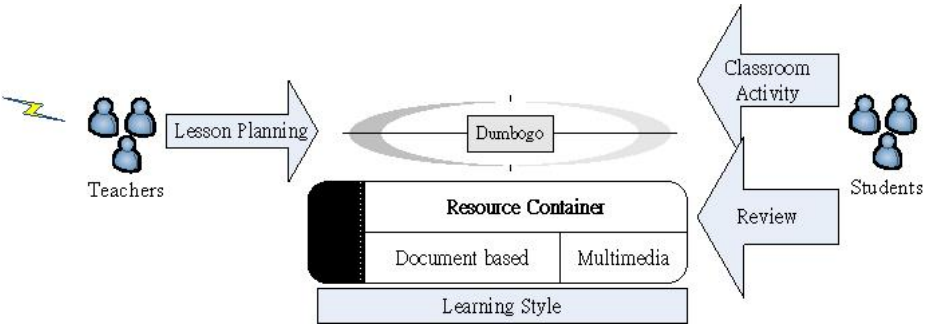


Fig. 5. Pedagogy Dumbogo

5 Conclusion and Future Work

We have been implementing our ideas to our Educational Cloud system [17] for data, knowledge sharing/updating between teachers and students, and have some preliminary results. Finally, we are also following the revolution of iPads, tablets, and panels, and we will improve our Dumbogo to make our Educational Cloud system work better by combining cloud computing and mobile computing.

In the future, we will develop a unified communication and collaboration (uc&c) environment which will simplify installations of low-cost educational clouds by schools and teachers around the world. The cloud's ontology and apps will support blended learning as well as support the organization of knowledge exchange clubs among students and teachers. Students will be encouraged by using a standardized certification of student levels and granting of special privileges to qualified students. Documents and educational resources that are developed by students and teachers will be retrievable by using parallel, distributed searches within a configuration for each knowledge exchange group.

Acknowledgments

We would like to thank the sponsor by HP Innovation in education grants in 2010.

References

1. http://win.dgbas.gov.tw/dgbas03/bs7/calendar/calendar.asp?
2. http://en.wikipedia.org/wiki/M-shape_Society
3. Grossman, R.L.: The Case for Cloud Computing. IT Professional 11, 23–27 (2009)

4. Gaoliang, P., Gongdong, W., Wenjian, L., Haiquan, Y.: A desktop virtual reality-based interactive modular fixture configuration design system vol. 42, pp. 432–444 (2010)
5. http://code.google.com/intl/zh-TW/webtoolkit/
6. Phipps, R., Merisotis, J.: What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education. The Institute for Higher Education Policy, Washington, DC (1999)
7. Yang, C.-Y., Chang, C.-T., Chien, L.-R.: A teacher control model of blended learning environment based on a typical classroom in secondary school. In: 2010 2nd International Conference on Computer Engineering and Technology, Chendu, China, vol. 7, pp. 720–724 (2010)
8. Osguthorpe, R.T., Graham, C.R.: Blended learning environments: Definitions and directions. The Quarterly Review of Distance Education 4(3), 227–233 (2003)
9. Chung, J., Davis, I.K.: An instructional theory for learner control: Revisited. In: Proceedings of the Annual National Convention of the Association for Educational Communications and Technology, pp. 72–86 (1995)
10. Bielawski, L., Metcalf, D.: Blended learning: integrating knowledge, performance support, and online learning. HRD Press, Amherst (2003)
11. DET: Blended Learning, NSW Department of Education and Training (2003)
12. Snow, R.E.: Individual Difference in the Design of Educational Programs. American Psychologist 41(10), 1029–1039 (1986)
13. Kolb, D.A., Rubin, I.M., McIntyre, J.M.: Organizational Psychology. A book of Reading (1979)
14. Kolb, A.Y., Kolb, D.A.: The Kolb's Learning Style Inventory-Version 3.1 2005 Technical Specifications. Hay Resource Direct, Boston (2005)
15. http://en.wikipedia.org/wiki/Dumbo
16. http://en.wikipedia.org/wiki/Hadoop#Hadoop_Distributed_File_System
17. Buehrer, D.J., Chien, L.-R.: The Educational Cloud, http://www.hp.com/hpinfo/newsroom/press/2010/100706xa.html

Author Index