

Yinglin Wang
Tianrui Li (Eds.)

Foundations of Intelligent Systems

 Springer

Advances in Intelligent and Soft Computing

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 107. P. Melo-Pinto, P. Couto, C. Serôdio,
J. Fodor, and B. De Baets (Eds.)
Eurofuse 2011, 2011
ISBN 978-3-642-24000-3

Vol. 108. Y. Wang (Ed.)
Education and Educational Technology, 2011
ISBN 978-3-642-24774-3

Vol. 109. Y. Wang (Ed.)
*Education Management, Education Theory
and Education Application, 2011*
ISBN 978-3-642-24771-2

Vol. 110. L. Jiang (Ed.)
*Proceedings of the 2011 International
Conference on Informatics, Cybernetics,
and Computer Engineering (ICCE 2011)
November 19-20, 2011, Melbourne,
Australia, 2011*
ISBN 978-3-642-25184-9

Vol. 111. L. Jiang (Ed.)
*Proceedings of the 2011 International
Conference on Informatics, Cybernetics, and
Computer Engineering (ICCE 2011) November
19-20, 2011, Melbourne, Australia, 2011*
ISBN 978-3-642-25187-0

Vol. 112. L. Jiang (Ed.)
*Proceedings of the 2011 International
Conference on Informatics, Cybernetics, and
Computer Engineering (ICCE 2011) November
19-20, 2011, Melbourne, Australia, 2011*
ISBN 978-3-642-25193-1

Vol. 113. J. Altmann, U. Baumöl, and
B.J. Krämer (Eds.)
Advances in Collective Intelligence 2011, 2011
ISBN 978-3-642-25320-1

Vol. 114. Y. Wu (Ed.)
*Software Engineering and Knowledge
Engineering: Theory and Practice, 2011*
ISBN 978-3-642-03717-7

Vol. 115. Y. Wu (Ed.)
*Software Engineering and Knowledge
Engineering: Theory and Practice, 2011*
ISBN 978-3-642-03717-7

Vol. 116. Yanwen Wu (Ed.)
*Advanced Technology in Teaching - Proceedings
of the 2009 3rd International Conference on
Teaching and Computational Science
(WTCS 2009), 2012*
ISBN 978-3-642-11275-1

Vol. 117. Yanwen Wu (Ed.)
*Advanced Technology in Teaching - Proceedings
of the 2009 3rd International Conference on
Teaching and Computational Science
(WTCS 2009), 2012*
ISBN 978-3-642-25436-9

Vol. 118. A. Kapczynski, E. Tkacz,
and M. Rostanski (Eds.)
*Internet - Technical Developments and
Applications 2, 2011*
ISBN 978-3-642-25354-6

Vol. 119. Tianbiao Zhang (Ed.)
*Future Computer, Communication, Control
and Automation, 2011*
ISBN 978-3-642-25537-3

Vol. 120. Nicolas Loménie, Daniel Racoceanu, and
Alexandre Gouaillard (Eds.)
*Advances in Bio-Imaging: From Physics to Signal
Understanding Issues, 2011*
ISBN 978-3-642-25546-5

Vol. 121. Tomasz Traczyk and
Mariusz Kaleta (Eds.)
*Modeling Multi-commodity Trade: Information
Exchange Methods, 2011*
ISBN 978-3-642-25648-6

Vol. 122. Yinglin Wang and Tianrui Li (Eds.)
Foundations of Intelligent Systems, 2011
ISBN 978-3-642-25663-9

Yinglin Wang and Tianrui Li (Eds.)

Foundations of Intelligent Systems

Proceedings of the Sixth International
Conference on Intelligent Systems and
Knowledge Engineering, Shanghai, China,
Dec 2011 (ISKE2011)



Springer

Editors

Prof. Yinglin Wang
Department of Computer Science
and Engineering
Shanghai Jiao Tong University
800 Dongchuan Road, Shanghai 200240,
China
E-mail: ylwang@sjtu.edu.cn

Prof. Tianrui Li
School of Information Science
and Technology
Southwest Jiaotong University
Chengdu, Sichuan Province, 610031,
China
E-mail: trli@swjtu.edu.cn

ISBN 978-3-642-25663-9

e-ISBN 978-3-642-25664-6

DOI 10.1007/978-3-642-25664-6

Advances in Intelligent and Soft Computing

ISSN 1867-5662

Library of Congress Control Number: 2011942150

© 2011 Springer-Verlag Berlin Heidelberg

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

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typeset by Scientific Publishing Services Pvt. Ltd., Chennai, India

Printed on acid-free paper

5 4 3 2 1 0

springer.com

Preface

We would like to extend our warmest welcome to each conference attendee. The 2011 International Conference on Intelligent Systems and Knowledge Engineering (ISKE2011) is the sixth in a series of ISKE conferences, which follows the successful ISKE2006 in Shanghai, ISKE2007 in Chengdu, and ISKE2008 in Xiamen, China, ISKE2009 in Hasselt, Belgium, and ISKE2010 in Hangzhou, China. ISKE2011 will be held in Shanghai, China, during December 15–17, 2011. It has been our pleasure as Program Committee Co-Chairs and Conference Co-Chair to organize this impressive scientific and technical program and the technical proceedings. ISKE2011 emphasizes current practices, experiences and promising new ideas in the broad area of intelligent systems and knowledge engineering. It provides a forum for researchers and practitioners around the world to present their latest results in research and applications and exchange new ideas in this field. ISKE 2011 is technically organized by Shanghai Jiao Tong University, and co-sponsored by California State University, Southwest Jiaotong University, Belgian Nuclear Research Centre (SCK•CEN).

We received 605 submissions from 26 countries and regions. We are very pleased with this level of submission and international participation. From these 605 submissions, the program committee selected 262 papers (including 109 full papers and 153 short papers), based on their originality, significance, correctness, relevance, and clarity of presentation, to be included in the proceedings. The acceptance rate of full papers is 18%, which we are proud of. The acceptance rate of short papers is 25%. Besides the papers in the conference proceedings, we also selected 44 papers from the submissions to be published in the Journal of Shanghai Jiao Tong University and the Journal of Donghua University. All the accepted papers will be presented or posted at the conference. Each of them was reviewed by two or more reviewers and the authors were asked to address each comment made by the reviewers for improving the quality of their papers. The acceptance rate of all the papers in the proceedings is 43%.

The accepted papers in the proceedings are contained in three volumes respectively based on the topics of the papers. The proceedings include “Volume I: Foundations of Intelligent Systems”, “Volume II: Knowledge Engineering and Management” and “Volume III: Practical Applications of Intelligent Systems”. Topics covered by the accepted papers in each volume of the proceedings are as follows:

Volume 1: Foundations of Intelligent Systems

Artificial Intelligence	46
Pattern Recognition, Image and Video Processing	40
Cognitive Science and Brain-Computer Interface	1

Volume 2: Knowledge Engineering and Management

Optimization and Biological Inspired Computation	12
Distributed and Parallel Intelligence	6
Robotics	11
Knowledge Engineering and Management	13
Data Mining, NLP and Information Retrieval	23
Data Simulation and Information Integration	11
Formal Engineering & Reliability	10

Volume 3: Practical Applications of Intelligent Systems

Social Computing, Mobile and Service Computing	8
Intelligent Game and Human Computer Interaction	6
Intelligent Engineering Systems	46
Intelligent Control Systems	9
Intelligent GIS, Networks or the Internet of Things	5
Social Issues of Knowledge Engineering	15

Accepted papers come from 23 countries, which shows that ISKE 2011 is a well-represented major international event, and their statistics (only papers of the proceeding, not include 44 papers which will be published in two journals) in terms of country are as follows:

China	212	Iran	2
Spain	8	Belgium	2
Australia	7	France	1
Brazil	5	Serbia	1
UK	6	Japan	1
USA	4	Vietnam	1
Russia	3	Sweden	1
Finland	3	Germany	1
Turkey	3	Saudi Arabia	1
Canada	2	Pakistan	1
Algeria	2	Korea	1
Poland	2		

ISKE 2011 consists of a three-day conference which includes paper and poster tracks, three invited keynote talks and two tutorials. The keynotes, tutorials and technical sessions cover a wide range of topics in intelligent systems and knowledge engineering.

The three invited speakers are Witold Pedrycz, University of Alberta, Canada; Ronald R. Yager, Iona College, New Rochelle, USA; and Zhi-Hua Zhou, Nanjing University, China. Witold Pedrycz will give a talk on granular models of time series and spatiotemporal data under the title of “User-Centric Models of Temporal and

Spatiotemporal Data: A Perspective of Granular Computing.” He will discuss a new category of models in which the mechanisms of description, processing, and predicting temporal and spatiotemporal data are expressed in the language of information granules, especially fuzzy sets and intervals. Ronald R. Yager’s talk is entitled “Intelligent Social Network Modeling.” He will discuss an approach to enrich the social network modeling by introducing ideas from fuzzy sets and related granular computing technologies. Zhi-hua Zhou will discuss the current research results of his group in the machine learning area.

The two invited tutorial speakers are Gio Wiederhold, Stanford University and Jie Lu, University of Technology, Sydney (UTS), Australia. Gio Wiederhold’s tutorial entitled “What is Your Software Worth?” will describe how the value of software can be estimated, and emphasize that awareness of the value of the product of one’s knowledge and effort can help in making decisions on the design and the degree of effort to be made. Jie Lu’s tutorial entitled “Personalized Recommender Systems for e-Government and e-Business Intelligence” will introduce several recommendation approaches, including case-based recommendation, ontology-based recommendation, fuzzy measure based recommendation, trust social networks-based recommendation related approaches and, in particular, present the recent developments made by her group in recommender systems and their applications in e-government and e-business intelligence.

As Program Committee Co-chairs and Conference Co-chair, we are grateful to all the authors who chose to contribute to ISKE2011. We want to express our sincere appreciation to the Program Committee Members listed below and to the additional reviewers for their great and quality work on reviewing and selecting the papers for the conference. We also would like to thank the webmasters, the registration secretary and financial secretary for their hard work. Last but certainly not the least, we would like to thank all the people involved in the organization and session-chairing of this conference. Without their contribution, it would not have been possible to produce this successful and wonderful conference. At this special occasion, we would especially like to acknowledge our respects and heartfelt gratitude to Professor Da Ruan, the Conference Co-chair of ISKE 2011 and the leading initiator of the ISKE conference series, for his hard work to prepare for this year’s conference. Professor Da Ruan worked tirelessly for the conference until he suddenly passed away on July 31. Our thoughts and prayers are with his family. Besides of the above, we also thank all the sponsors of the conference, the National Science Foundation of China (No. 60873108, No. 60773088) and the Springer Publishing Company for their support in publishing the proceedings of ISKE 2011.

Finally we hope that you find ISKE2011 programs rewarding and that you enjoy your stay in the beautiful city of Shanghai.

December 15–17, 2011

Tianrui Li
Program Committee Chair

Yinglin Wang
Program Committee Co-chair

Du Zhang
Conference Co-chair

Organizing Committee

Honorary Co-chairs

L.A. Zadeh	University of California, Berkeley, USA
Gio Wiederhold	Stanford University, USA
H.-J. Zimmermann	Aachen Institute of Technology, Germany
Etienne E. Kerre	Ghent University, Belgium

Conference Co-chairs

Da Ruan	Belgian Nuclear Research Centre, Belgium
Du Zhang	California State University, USA
Athman Bouguettaya	RMIT University, Australia
Javier Montero	Complutense University of Madrid, Spain
Fuchun Sun	Tsinghua University, China

Steering Committee Co-chairs

Ronald R. Yager	Iona College, New Rochelle, USA
Jyrki Nummenmaa	University of Tampere, Finland
Wensheng Zhang	Chinese Academy of Sciences, China
Weiming Shen	National Research Council of Canada, Canada
Koen Vanhoof	University of Hasselt, Belgium

Organization Chair

Yinglin Wang	Shanghai Jiao Tong University, China
--------------	--------------------------------------

Local Organization Co-chair

Hongming Cai	Shanghai Jiao Tong University, China
--------------	--------------------------------------

Program Chair

Tianrui Li Southwest Jiaotong University, China

Program Co-chairs

Yinglin Wang	Shanghai Jiao Tong University, China
Luis Martinez Lopez	University of Jaén, Spain
Hongtao Lu	Shanghai Jiao Tong University, China
Hongming Cai	Shanghai Jiao Tong University, China
Xuelong Li	Chinese Academy of Sciences, China

Publication Chair

Min Liu Tongji University, China

Special Session Co-chairs

Jie Lu	University of Technology, Sydney, Australia
Cengiz Kahraman	Istanbul Technical University, Turkey
Victoria Lopez	Complutense University of Madrid, Spain
Zheyang Zhang	University of Tampere, Finland

Poster Session Co-chairs

Guangquan Zhang	University of Technology, Sydney, Australia
Jun Liu	University of Ulster at Jordanstown, UK

Publicity Co-Chairs

Wujun Li	Shanghai Jiao Tong University, China
Xianyi Zeng	ENSAIT Textile Institute, France
Jiacun Wang	Monmouth University, USA
Michael Sheng	The University of Adelaide, Australia
Dacheng Tao	University of Technology, Sydney, Australia

Program Committee Members

Abdullah Al-Zoubi (Jordan)
Andrzej Skowron (Poland)
Athena Tocatlidou (Greece)
B. Bouchon-Meunier (France)
Benedetto Matarazzo (Italy)
Bo Yuan (USA)
Cengiz Kahraman (Turkey)
Chien-Chung Chan (USA)
Cornelis Chris (Belgium)
Dacheng Tao (Australia)
Davide Ciucci (Italy)
Davide Roverso (Norway)
Du Zhang (USA)
Enrico Zio (Italy)
Enrique Herrera-Viedma (Spain)
Erik Laes (Belgium)
Etienne E. Kerre (Belgium)
Francisco Chiclana (UK)
Francisco Herrera (Spain)
Fuchun Sun (China)
Gabriella Pasi (Italy)
Georg Peters (Germany)
Germano Resconi (Italy)
Guangquan Zhang (Australia)
Guangtao Xue (China)
Gulcin Buyukozkan (Turkey)
Guolong Chen (China)
Guoyin Wang (China)
H.-J. Zimmermann (Germany)
Hongjun Wang (China)
Hongming Cai (China)
Hongtao Lu (China)
I. Burhan Turksen (Canada)
Irina Perfilieva (Czech Republic)
Jan Komorowski (Sweden)
Janusz Kacprzyk (Poland)
Javier Montero (Spain)
Jer-Guang Hsieh (Taiwan, China)
Jesús Vega (Spain)
Jiacun Wang (USA)
Jianbo Yang (UK)
Jie Lu (Australia)
Jingcheng Wang (China)
Jitender S. Deogun (USA)
Jouni Jarvinen (Finland)
Juan-Carlos Cubero (Spain)
Jun Liu (UK)
Jyrki Nummenmaa (Finland)
Koen Vanhoof (Belgium)
Krassimir Markov (Bulgaria)
Liliane Santos Machado (Brasil)
Lisheng Hu (China)
Luis Magdalena (Spain)
Luis Martinez López (Spain)
Lusine Mkrtchyan (Italy)
Madan M. Gupta (Canada)
Martine De Cock (Belgium)
Masoud Nikravesch (USA)
Michael Sheng (Australia)
Mihir K. Chakraborty (India)
Mike Nachtegaal (Belgium)
Mikhail Moshkov (Russia)
Min Liu (China)
Peijun Guo (Japan)
Pierre Kunsch (Belgium)
Qi Wang (China)
Qingsheng Ren (China)
Rafael Bello (Cuba)
Richard Jensen (UK)
Ronald R. Yager (USA)
Ronei Marcos de Moraes (Brasil)
Ryszard Janicki (Canada)
S. K. Michael Wong (Canada)
Shaojie Qiao (China)
Shaozi Li (China)
Sheela Ramanna (Canada)
Su-Cheng Haw (Malaysia)
Suman Rao (India)
Sushmita Mitra (India)
Takehisa Onisawa (Japan)
Tetsuya Murai (Japan)
Tianrui Li (China)
Tzung-Pei Hong (Taiwan, China)
Ufuk Cebeci (Turkey)
Victoria Lopez (Spain)
Vilem Novak (Czech Republic)
Weiming Shen (Canada)
Weixing Zhu (China)

Wensheng Zhang (China)
Witold Pedrycz (Canada)
Wujun Li (China)
Xianyi Zeng (France)
Xiaogang Jin (China)
Xiaoqiang Lu (China)
Xiao-Zhi Gao (Finland)
Xuelong Li (China)
Xun Gong (China)
Yan Yang (China)
Yangguang Liu (China)
Yanmin Zhu (China)
Yaochu Jin (Germany)

Yasuo Kudo (Japan)
Yi Tang (China)
Yinglin Wang (China)
Yiyu Yao (Canada)
Yongjun Shen (Belgium)
Yuancheng Huang (China)
Zbigniew Suraj (Poland)
Zbigniew W. Ras (USA)
Zheyang Zhang (Finland)
Zhong Li (Germany)
Zhongjun He (China)
Zhongzhi Shi (China)

Additional Reviewers

Jianmei Guo
Peng Li
Xin Xu
Heming Xu

Hongbin Yu
Xianzhong Long
Yangcheng He

Volunteers

Registration Secretary
Financial Secretary
Web Masters
Conference Assistant

Shanshan Feng
Genzhen Chen
Dazhi Li and Guangxin Wang
Jinsong Zhang, Liang Tao, Jian Li,
Jinwei Pang, Jun Shi, Ruixin Zhang,
Yi Huang, Minglu Zhang, and Cai Chen

Sponsors



Shanghai JiaoTong University, China



The California State University, USA



Southwest Jiaotong University, China



Belgian Nuclear Research Centre, Belgian

Contents

Invited Talks

**User-Centric Models of Temporal and Spatiotemporal Data:
A Perspective of Granular Computing** XXIII
Witold Pedrycz

Intelligent Social Network Modeling XXV
Ronald R. Yager

What is Your Software Worth? XXVII
Gio Wiederhold

**Personalized Recommender Systems for e-Government and e-Business
Intelligence** XXIX
Jie Lu

Part I: Artificial Intelligence

**Efficient Closed Iterative Patterns Mining Algorithm via Prime-Block
Encoding** 3
Zhixin Ma, Zhe Ding, Yusheng Xu

Pool-Based Active Learning with Query Construction 13
Shanhong Zhang, Jian Yin, Weizhao Guo

**The Research of Soft Measurement Method Based on Sintering Process
Permeability Index** 23
Jinyu Teng, Xiaoxin Zhang

**A Neighborhood Preserving Based Semi-supervised Dimensionality
Reduction Method for Cancer Classification** 29
Xianfa Cai, Jia Wei, Guihua Wen, Jie Li

A Layered Model of Artificial Emotion Merging with Attitude	39
<i>Qijun Luo, Ang Zhao, Hongxiang Zhang</i>	
Fast Extraction Strategy of Support Vector Machines	49
<i>Wei Wu, Qiang Yang, Wenjun Yan</i>	
A Local Feature Selection Approach for Clustering	55
<i>Bing Gui</i>	
Semantics-Preserving Fusion of Structures of Probabilistic Graphical Models	63
<i>Kun Yue, Yunlei Zhu, Kailin Tian, Weiyi Liu</i>	
A Hybrid Genetic Algorithm for Scheduling and Selecting a Project Portfolio	69
<i>Bo Shi, Hong Wang, Lu Qi</i>	
Tibetan Processing Key Technology Research for Smart Mobile Phone Based on Symbian	77
<i>Nyima Trashi, Qun Nuo, Yong Tso, Pu Dun, Gama Zhaxi, Niluo Qiongda</i>	
PCNN Automatic Parameters Determination in Image Segmentation Based on the Analysis of Neuron Firing Time	85
<i>Xiangyu Deng, Yide Ma</i>	
Messy Genetic Algorithm for the Optimum Solution Search of the HTN Planning	93
<i>Jiangfeng Luo, Cheng Zhu, Weiming Zhang</i>	
Research on Rule-Based Reasoning Methods Oriented on Information Resource Ontology	99
<i>Gang Liu, Lifu Feng, Ying Liu, Zheng Wang</i>	
Generating Descriptions of Incomplete City-Traffic States with Agents	105
<i>Grzegorz Popek, Ryszard Kowalczyk, Radoslaw P. Katarzyniak</i>	
Target Factor Algorithm of PSC Ship-Selecting System Based on Rough Set Theory and Hierarchic Analysis	115
<i>Zhonghua Sun, Tingting Yang</i>	
Route Planning Based on Combination of Artificial Immune Algorithm and Ant Colony Algorithm	121
<i>Qingfeng Wang and Yuhui Wang</i>	
Hyperspectral Image Classification Using Support Vector Machines with an Efficient Principal Component Analysis Scheme	131
<i>Pinliang Dong, Jianguo Liu</i>	

A Practicable Heuristic Attributes Reduction Algorithm for Ordered Information Systems	141
<i>Wei Li, Feng Liu, Zhi-hong Zhao</i>	
An Improved Abrams-Strogatz Model Based Protocol for Agent Competition and Strategy Designing	151
<i>Cunhua Li, Yun Hu, Lanlan Sun</i>	
α-Quasi-Lock Semantic Resolution Method for Linguistic Truth-Valued Lattice-Valued Propositional Logic $\mathcal{L}_{V(n \times 2)}^P(X)$	159
<i>Xiaomei Zhong, Jun Liu, Shuwei Chen, Yang Xu</i>	
A Real-Time Software GPS Receiver Based on MAX2769	171
<i>Xiao Xu, Yijin Chen</i>	
Duality in Lattice Implication Algebra	181
<i>Li Zhao, Yang Xu</i>	
Kernel Construction via Generalized Eigenvector Decomposition	191
<i>Yong Liu, Shizhong Liao</i>	
Probabilistic Model Combination for Support Vector Machine Using Positive-Definite Kernel-Based Regularization Path	201
<i>Ning Zhao, Zhihui Zhao, Shizhong Liao</i>	
The Relationship of Filters in Lattice Implication Algebra	207
<i>Ruijuan Lv, Yang Xu</i>	
Effects of Vision Field on Evolution of Cooperation among Mobile Agents	219
<i>Wei-Ye Wang, Zhou Zhou, Xiao-Long Jiang</i>	
Deviating from Common Context in Individual Semiosis in Multi-Agent Systems	229
<i>Wojciech Lorkiewicz, Radoslaw Katarzyniak, Ryszard Kowalczyk</i>	
Active Discovery Based Query Federation over the Web of Linked Data	239
<i>Xuejin Li, Zhendong Niu, Chunxia Zhang</i>	
A Data Imputation Method with Support Vector Machines for Activity-Based Transportation Models	249
<i>Banghua Yang, Davy Janssens, Da Ruan, Mario Cools, Tom Bellemans, Geert Wets</i>	
The Study on Integer Overflow Vulnerability Detection in Binary Executables Based upon Genetic Algorithm	259
<i>Baojiang Cui, Xiaobing Liang, Jianxin Wang</i>	

Case-Based Reasoning Genetic Algorithm for Rectangle and Circle Packing Problem with Equilibrium Constraints	267
<i>Ziqiang Li, Meng-juan Dong</i>	
Feature Discriminability for Pattern Classification Based on Neural Incremental Attribute Learning	275
<i>Ting Wang, Sheng-Uei Guan, Fei Liu</i>	
Selective Ensemble Approach for Classification of Datasets with Incomplete Values	281
<i>Yan Wang, Yi Gao, Ruimin Shen, Fan Yang</i>	
Hesitant Fuzzy Linguistic Term Sets	287
<i>Rosa M. Rodríguez, Luis Martínez, Francisco Herrera</i>	
A Conceptual Model for Risk-Based Situation Awareness	297
<i>Mohsen Naderpour, Jie Lu, Etienne Kerre</i>	
On Partial Comparability and Fuzzy Preference-Aversion Models	307
<i>Camilo Franco, Javier Montero, J. Tinguaro Rodríguez</i>	
Honesty-Rate Measurement: A Novel Approach for the Fragile Trust Inside the DIDS	317
<i>Peijian Chen, Yuexiang Yang, Hailong Wang, Chuan Tang, Jie He</i>	
Using Cooperative Clustering to Solve Multiclass Problems	327
<i>Chuanhuan Yin, Shaomin Mu, Shengfeng Tian</i>	
Total Colorings of Planar Graphs with Maximum Degree Seven and without 3-Cycles Adjacent to 5-Cycles	335
<i>Guangde Liu, Bing Wang, Jian-liang Wu</i>	
A Meta-Strategy for Coordinating of One-to-Many Negotiation over Multiple Issues	343
<i>Khalid Mansour, Ryszard Kowalczyk</i>	
Further Research of Generated Filters in Lattice Implication Algebra	355
<i>Ling Guo, Yang Xu, Shaokun Du</i>	
Dynamic Task Allocation and Action Coordination under Uncertain Environment	365
<i>Chengli Liu, Wei Zeng, Hongtao Zhou, Lei Cao, Yang Yang</i>	
Representation and Acquisition of Feature Value of Learner Model in Adaptive Learning System	371
<i>Bing Jia, Yongjian Yang, Jun Zhang</i>	
MP-IR: A Market-Oriented Mobile Agents System for Distributed Information Retrieval	379
<i>Djamel Eddine Menacer, Habiba Drias, Christophe Sibertin-Blanc</i>	

Framework for Goal-Driven Negotiation Process	391
<i>Ying Lei</i>	
Inconsistency in Multi-Agent Systems	401
<i>Du Zhang</i>	
Part II: Pattern Recognition, Image and Video Processing	
Emotional Speech Recognition Based on Syllable Distribution Feature Extraction	415
<i>Haiying Zhang</i>	
Face Recognition Based on the Second-Generation Curvelet Transform Domain and KPCA	421
<i>Xian Wang, Xin Mu, Yan Zhang, Fangsheng Zhang</i>	
Marine Spill Oil SAR Image Segmentation Based on Maximum Entropy and CV Model	427
<i>Yang Ji, Yiquan Wu, Yi Shen</i>	
Multilevel Thresholding Based on Exponent Gray Entropy and Niche Chaotic Particle Swarm Optimization	437
<i>Yi Shen, Yiquan Wu, Yang Ji</i>	
Learning Bag-of-Words Models Using Sparse Partial Least Squares	445
<i>Jingneng Liu, Guihua Zeng</i>	
Blind Watermark Algorithm Based on QR Barcode	457
<i>Meifeng Gao, Bing Sun</i>	
Cloth Pattern Simulation Based on a 1/f Noise Method	463
<i>Beibei Li, Zhihong Zhao</i>	
An Anti-noise Determination on Fractal Dimension for Digital Images	469
<i>Ying Shi, Shu Cheng, Shuhai Quan, Ting Bai</i>	
A Spectral Matching for Shape Retrieval Using Pairwise Critical Points	475
<i>Zhen Pan, Guoqiang Xiao, Kai Chen, Zhenghao Li</i>	
Skew Detection of Fabric Images Based on Edge Detection and Projection Profile Analysis	483
<i>Zhoufeng Liu, Jie Huang, Chunlei Li</i>	
Improved Algorithm of Separation and Identification of Touching Kernels and Foreign Materials in Digital Images	489
<i>Zhining Liu, Lei Yan</i>	
Object Detection Based on Multiclass Discriminative Field	495
<i>Xiaofeng Zhang, Qiaoyu Sun, Yue Lu</i>	

Co-processing Method for Automotive Vibration Signals on JBeam	501
<i>Guofeng Qin, Minhu Fan, Qiyan Li</i>	
A Method of Image Segmentation Based on Improved Adaptive Genetic Algorithm	507
<i>Wenjiao Yu, Mengxing Huang, Donghai Zhu, Xuegang Li</i>	
Resistance Identifier Recognition Based on Wavelet Transform and LBP Operator	517
<i>Chong-quan Zhong, Yan-dong Zhu</i>	
Image Processing Methods for V-Shape Weld Seam Based on Laser Structured Light	527
<i>Tao Qin, Ke Zhang, Jingyu Deng, Xin Jin</i>	
Melon Image Segmentation Based on Prior Shape LCV Model	537
<i>Yubin Miao, Qiang Zhu</i>	
An Improved Method for Terrain Mapping from Descent Images	547
<i>Xiaoliang Xue, Cai Meng, Yang Jia</i>	
Best View Selection of 3D Object Based on Sample Learning	557
<i>Zhi Liu, Yipan Feng, Qihua Chen, Xiang Pan</i>	
An Improved Retina Modeling for Varying Lighting Face Recognition	563
<i>Yong Cheng, Yingkun Hou, Zuoyong Li</i>	
Automated Object Length Measurement Applied to AFM, STM and TEM Images Based on the Snake Model	569
<i>Leandro Marturelli, Lilian Costa, Geraldo Cidade</i>	
Automatic Smudge Cell Recognition Based on Unimodal Region and Weak Edge Features	575
<i>Guohui Qiao, Minglei Sun, Guanghua Zong, Fanggu Wu, Suling Huang, Shichuan Tang</i>	
Recognition of Touching Erythrocytes via Contour Radial Uniformity and Sector Region Distribution Features	581
<i>Minglei Sun, Di Wang, Wen Wen, Rong Zhang, Shichuan Tang, Bin Zhang</i>	
A Continuation Log-Barrier Method for ℓ_1-regularized Least Square	591
<i>Min Zhang, Dongfang Chen</i>	
Use of Imaging Techniques to Obtain 3D Models of Small Insects	603
<i>Franxavier Centeno, Ángela López Benítez, Carles Domènech, Francesc Pérez-Ràfols, Joaquim Lloveras Macià</i>	

Affine Object Tracking Using Kernel-Based Region Covariance Descriptors	613
<i>Bo Ma, Yuwei Wu, Fengyan Sun</i>	
A Pulmonary Nodules Detection Method Using 3D Template Matching	625
<i>Ting Gao, Xiwen Sun, Yuanjun Wang, Shengdong Nie</i>	
An Improved Iterative Binary Coloring Procedure for Color Image Segmentation	635
<i>Javier Montero, Susana Muñoz, Daniel Gómez</i>	
Moving Objects Detection Using Adaptive Region-Based Background Model in Dynamic Scenes	641
<i>Lin Gao, Yong Fan, Niannian Chen, Yufeng Li, Xiaorong Li</i>	
New Shape-from-Shading Method with Near-Scene Point Lighting Source Condition	653
<i>Lei Yang, Shiwei Ma, Bo Tian</i>	
Mixture of Subspace Learning with Adaptive Dimensionality: A Self-Organizing Approach	665
<i>Huicheng Zheng</i>	
Dialect Identification Based on S-Transform and Singular Value Decomposition	673
<i>Yan He, Fengqin Yu</i>	
Local Structure Recognition of Point Cloud Using Sparse Representation	679
<i>Pei Luo, Zhuangzhi Wu, Teng Ma</i>	
Pyroelectric Infrared Sensors for Human Identification Using Non-negative Matrix Factorization and BP Neural Network	685
<i>Ning Zhao, Fangmin Li, Sheng Wang</i>	
Bimodal Emotion Recognition Based on Speech Signals and Facial Expression	691
<i>Binbin Tu, Fengqin Yu</i>	
Memory-Based Multi-camera Handover with Non-overlapping Fields of View	697
<i>Xiaoyan Sun, Faliang Chang, Jiangbao Li</i>	
The Face Recognition Algorithm Based on Curvelet Transform and CSVD	705
<i>Shulin Song, Yan Zhang, Xian Wang, Xin Mu</i>	
Background Modeling for Fire Detection	715
<i>Yan Yang, Xiaopeng Hu, Chaofei Zhang, Yi Sun</i>	

Research on the Technology of Video Semantic Retrieval Based on Structured Semantic Strings	721
<i>Jing Yuan, Quan Zheng, Zhijun Sun, Song Wang</i>	
Color Image Segmentation Algorithm Based on Affinity Propagation Clustering	731
<i>Lei Wang, Lin Zhang</i>	
Part III: Cognitive Science and Brain-Computer Interface	
A EEG-Based Brain Computer Interface System towards Applicable Vigilance Monitoring	743
<i>Hongfei Ji, Jie Li, Lei Cao, Daming Wang</i>	
Author Index	751

User-Centric Models of Temporal and Spatiotemporal Data: A Perspective of Granular Computing

Witold Pedrycz

Department of Electrical & Computer Engineering
University of Alberta, Edmonton Canada
and
Systems Research Institute, Polish Academy of Sciences
Warsaw, Poland
pedrycz@ee.ualberta.ca

Abstract

One of the ultimate objectives of intelligent data analysis is to develop models of data that are user-centric. The human centrality of such pursuits means that a process of analysis along with the obtained results are made transparent to the user and come with a significant degree of flexibility, which helps achieve a sound tradeoff between accuracy and interpretability of results. The perception of data, as realized by humans, inherently invokes information granules (realized through numerous formal approaches including fuzzy sets, interval analysis, and rough sets) and their further processing. This helps establish a suitable level of abstraction at which the data are perceived, analyzed and their models are being formed. By casting the problem in the setting of Granular Computing, we develop a new category of models in which the mechanisms of description, processing, and predicting temporal and spatiotemporal data are expressed in the language of information granules, especially fuzzy sets and intervals.

In this talk, we show how a principle of justifiable information granularity leads to the realization of granular models of time series in which a construction of information granules is viewed as a certain optimization problem.

With regard to spatiotemporal data where their temporal facet as well as their spatial characteristics play a pivotal role, it is demonstrated how information granules are formed through an augmented collaborative clustering. The grouping is completed in the temporal and spatial domain in such a way an identity of relationships present in these two domains is retained. An auxiliary mechanism of information granulation is developed through an optimization of relational constraints (granular codebook) realized through a collection of information granules.

“(The full content will be available during the conference)”

Intelligent Social Network Modeling

Ronald R. Yager

Machine Intelligence Institute, Iona College
New Rochelle, NY 10801
yager@panix.com

Abstract

Web 2.0 has provided for a rapid growth of computer mediated social networks. Social relational networks are becoming an important technology in human behavioral modeling. Our goal here is to enrich the domain of social network modeling by introducing ideas from fuzzy sets and related granular computing technologies. We approach this extension in a number of ways. One is with the introduction of fuzzy graphs representing the networks. This allows a generalization of the types of connection between nodes in a network from simply connected or not to weighted or fuzzy connections. A second and perhaps more interesting extension is the use of Zadeh's fuzzy set based paradigm of computing with words to provide a bridge between a human network analyst's linguistic description of social network concepts and the formal model of the network. Another useful extension we discuss is vector-valued nodes. Here we associate with each node a vector whose components are the attribute values of the node. Using the idea of computing with words we are then able to intelligently query the network with questions that involve both attributes and connections. We see this as a kind of social network database theory. We shall look at some dynamic structures of network particularly the small worlds network.

“(The full content will be available during the conference)”

What Is Your Software Worth?

Gio Wiederhold

Professor Emeritus, Stanford University and MITRE Corporation
gio@cs.stanford.edu

Abstract

Much has been written about the cost of producing software, but that literature largely ignores the benefits of using that software. While software creators believe that their products are valuable, they are rarely called upon to quantify its benefits. Evaluation of software and its benefits in commerce is left to lawyers, economists, software vendors, or promoters. The results are often inconsistent.

This tutorial describes how the value of software can be estimated. The problem being addressed is that the value of software is essentially independent of the cost and effort spent to create it. A few brilliant lines of code can have a very high value, whereas a million lines of code that generate a report that nobody uses have little value. Awareness of the value of the product of one's knowledge and effort can help in making decisions on the design and the degree of effort to be made.

The tutorial will survey methods for valuing software based on the income it can generate. A principal approach is based on software growth, caused by needed maintenance. The valuation is with the accepted framework for valuing intellectual property (IP) in general.

My paper on that topic appeared in the Communications of the ACM, September 2006, but could not cover all of the issues. More material is available at <http://infolab.stanford.edu/pub/gio/inprogress.html#worth>. Software valuation is also covered in a course at Stanford University, CS207, <https://cs.stanford.edu/wiki/cs207/> Participants in the tutorial are encouraged to read the available information and engage in discussion of this challenging topic.

“(The full content will be available during the conference)”

Personalized Recommender Systems for e-Government and e-Business Intelligence

Jie Lu

Decision Systems & e-Service Intelligence Lab
Centre for Quantum Computation & Intelligent Systems
School of Software
Faculty of Engineering and Information Technology
University of Technology, Sydney
P.O. Box 123, Broadway, NSW 2007, Australia
jielu@it.uts.edu.au

Abstract

Web personalisation is an interdisciplinary topic that has been discussed in the literature about information systems, web intelligence, customer relationship management and marketing. Web personalisation is defined as any set of actions that tailor the web experience to a specific user or set of users, anticipating user needs to provide them with what they want or require without having to ask for it explicitly. A number of e-business and e-government development stage models have been proposed in the literature that focuses on classifying functions and features offered by current e-business and e-government. Most of these models have a common final stage which concentrates on providing fully integrated and personalised e-services for their constituents. Recommender systems have gained considerable attention in recent years and are the most successful implementation of web personalisation. Recommender systems use justifications to generate recommended products or services to customers and to ensure the customers like these products or services. These justifications can be obtained either from preferences directly expressed by customers, or induced, using data representing the customer experience. Recommender systems are achieving widespread success and have attracted researchers' attention in the field of e-business and e-government applications.

Recommender systems use different types of information filtering techniques to automatically identify and predict a set of interesting items on behalf of the users according to their personal preferences. The most notable classes of recommender system approaches include: (1) Content-based filtering--mainly depends on items' descriptions to generate personalised recommendations; (2) Collaborative Filtering (CF)--mainly depends on users ratings of items in a given domain, and works by computing the similarities between the profiles of several users on the basis of their provided ratings and generates new recommendations based on comparisons of user ratings; (3) Knowledge-based filtering--suggests items based on logical inferences about a user's needs and preferences; (4) Semantic-based filtering--exploits the semantic information associated with user and item descriptions to generate recommendations; (5) Trust-based filtering--exploits the level of trust between users in a social trust network and uses that knowledge to generate trustworthy recommendations; (6)

Hybrids-based filtering--combines two or more recommendation approaches to exploit their strengths and reduce their weaknesses.

This tutorial will introduce these recommendation approaches and, in particular, present the recent developments made by our Decision Systems and e-Service Intelligence (DeSI) lab in recommender systems and their applications in e-government and e-business intelligence , including case-based recommendation, ontology-based recommendation, fuzzy measure based recommendation, trust social networks-based recommendation related approaches and their applications in telecom companies and government-to-business services.

“(The full content will be available during the conference)”

Part I
Artificial Intelligence

Efficient Closed Iterative Patterns Mining Algorithm via Prime-Block Encoding^{*}

Zhixin Ma, Zhe Ding^{**}, and Yusheng Xu

School of Information Science and Engineering, Lanzhou University, Lanzhou, 730000, China
{mazhx, xuyusheng}@lzu.edu.cn, dingzhe0301@yahoo.com.cn

Abstract. In this paper, a novel algorithm which is called CIPMPBE (Closed Iterative Pattern Miner via Prime-Block Encoding) is proposed to mine closed iterative patterns. CIPMPBE is composed of three separated steps. In the first step, the positional information of all frequent i -sequences is generated. In the second step, the positional information of all frequent i -sequences and all instances of frequent $(i-1)$ -iterative patterns are used to obtain the positional information of all instances of frequent i -iterative patterns. In the third step, the positional information of all instances of frequent i -iterative pattern is used to obtain the positional information of the entire closed i -iterative pattern and get back to the first step to generate the positional information of closed $(i+1)$ -iterative pattern. For effective testing, a set of experiments were performed. The results of these experiments show that the time efficiency of CIPMPBE is better than that of CLIPER (CLOsed Iterative Pattern minER).

Keywords: Closed Iterative Pattern, Prime-Block Encoding.

1 Introduction

The task of sequential pattern mining which is proposed by Agrawal and Sriant is to discover temporal patterns that are supported by a significant number of sequences [1]. A pattern is supported by a sequence if it is a sub-sequence.

Iterative pattern mining is proposed by Lo *et al* in [3]. Iterative pattern is formed by a serious of events, and supported by a significant number of instances repeated within and across sequences. Iterative pattern mining which is similar to sequential pattern considers all sequences in the database rather than a single sequence. To reduce total number of iterative pattern without lose the information of iterative pattern, Lo *et al* proposed mining closed iterative pattern [3]. Iterative pattern mining has broad application such as prediction of software defect, feature discovery of software defect and software specification discovery, etc.

To improve the efficiency of iterative pattern mining and reduce time cost of the system, a novel algorithm called CIPMPBE is proposed in this paper. The algorithm is based on CLPER [3] and utilizes PRISM (PRIME-Encoding Based Sequence Mining) [4] to generate candidate and prune searching space.

^{*} The work was supported by the Fundamental Research Funds for Central Universities under grant No.lzugby-2010-91.

^{**} Corresponding author.

The rest of the paper is organized as follows: Section 2 discusses the related work of CIPMPBE. Section 3 introduces the basic concepts, methods of generating candidate and pruning strategies. Section 4 proposes CIPMPBE. A comprehensive experimental study is presented in Section 5. Finally, conclusions can be found in section 6.

2 Related Works

The sequential patterns mining problem is first proposed in [1] by Agrawal and Srikant and directed towards data analysis of market basket.

In [3], Lo *et al.* proposed iterative pattern mining. Iterative pattern mining is the extension of sequential pattern mining. The difference between sequential pattern mining and iterative pattern mining is that iterative pattern mines not only the pattern occurring to multiple sequences but also repeated within each sequence. Frequent iterative pattern is supported by a significant number of instances repeated within and across sequence.

In [4], Zaki *et al.* proposed an algorithm called PRISM. This algorithm is based on the novel notion of primal block encoding, which in turn is based on prime factorization theory.

Our previous work [5] were proposed two pruning strategies called DSEP (Dynamic Sequence Extension Pruning) and DIEP (Dynamic Item Extension Pruning) which can be used in all Apriori-like sequence mining algorithm or lattice-theoretic approach prune out infrequent candidate sequences during mining process.

3 Problems Statement

3.1 Basic Definitions and Concepts

Let I be a set of distinct events. A sequence $S = \langle k_1, k_2, \dots, k_m \rangle$ ($k_i \in I$) is an ordered list. The i th event in the sequence S is denoted as $S[i]$.

A pattern $P_1 = \langle a_1, a_2, \dots, a_n \rangle$ is considered as a subsequence of another pattern $P_2 = \langle b_1, b_2, \dots, b_m \rangle$ if there exist integers i_1, i_2, \dots, i_n such that $1 \leq i_1 < i_2 < \dots < i_n \leq m$ and $a_1 = b_{i_1}, a_2 = b_{i_2}, \dots, a_n = b_{i_n}$. The relationship is denoted by $P_1 \subseteq P_2$. The first item and last item of pattern P are denoted as $\text{first}(P)$ and $\text{last}(P)$ respectively. And the sequence database is denoted as SDB. The number of events in a sequence and pattern are respectively called the length of sequence and pattern.

Definition 1. Concatenation of two patterns $P_1(\langle a_1, \dots, a_n \rangle)$ and $P_2(\langle b_1, \dots, b_m \rangle)$ will generate a longer pattern $P_3(\langle a_1, \dots, a_n, b_1, \dots, b_m \rangle)$. Truncation operation is only applicable between a pattern and its suffix. Truncation of a pattern $P_3(\langle a_1, \dots, a_n, b_1, \dots, b_m \rangle)$ and a suffix $P_2(\langle b_1, \dots, b_m \rangle)$ will result in the pattern $P_1(\langle a_1, \dots, a_n \rangle)$. Patterns concatenation is denoted by ++, while pattern truncation is denoted by --.

Definition 2. Given a pattern $P(\langle P_1, \dots, P_n \rangle)$ and a string $S(\langle s_1, \dots, s_m \rangle)$, the erasure of S wrt. P , denoted by $\text{erasure}(S, P)$, is defined as a new string S_{erased} formed from S

where all events occurring in P are removed from S . Formally, S_{erased} is defined as $\langle se_1, se_2, \dots, se_k \rangle$ such that (1) $\forall i. se_i \notin P$ and (2) there exists a set of integers $\{i_1, i_2, \dots, i_k\}$ with $1 \leq i_1 < \dots < i_k \leq m$ and $se_{i_1} = s_{i_1}, \dots, se_{i_k} = s_{i_k}$ and $\forall j \notin \{i_1, i_2, \dots, i_k\}, s_j \in P$.

3.2 Semantic of Iterative Patterns

The origin of iterative pattern's definition is the common languages for specifying software behavioral requirement: Message Sequence Chart (MSC) (a standard of International Telecommunication Union (ITU)) and its extension, Live Sequence Chart (LSC). The iterative pattern instance definition capturing the total-ordering and one-to-one correspondence between events in the pattern and its instance can be expressed unambiguously in the form of Quantified Regular Expression (QRE). [3]

Definition 3. Given a pattern $P(p_1, p_2, \dots, p_n)$, a substring $S(s_1, s_2, \dots, s_m)$ of a sequence S in SDB is an instance of P if and only if it is of the following QRE expression

$$p_1 ; [-p_1, p_2, \dots, p_n]^* ; p_2 ; \dots ; [-p_1, p_2, \dots, p_n]^* ; p_n.$$

The symbols “;” “[−]” and “*” are denoted as concatenation operator, exclusion operator (i.e. $[-P, S]$ means any event except P and S), and standard kleene-star respectively.

An instance of iterative pattern is expressed as a triple $(S_{IDX}, I_{START}, I_{END})$ in this paper. S_{IDX} expresses the sequence index of a sequence S . I_{START} and I_{END} express the starting point and ending point of a substring in S . [3]

Definition 4. Given a pattern $P(p_1, \dots, p_n)$, a substring $ST(st_1, \dots, st_m)$ of a sequence S in SDB is an iterative pattern instance of P if and only if ① $\text{first}(P) = \text{first}(ST)$, ② $\text{last}(P) = \text{last}(ST)$ and ③ the following erasure constraint holds: $\text{erasure}(ST, \text{erasure}(ST, P)) = P$.

In this paper, the term “pattern instance” and the term “pattern” are interchanged respectively with the term “iterative pattern instance” and the term “iterative pattern”. The support of a pattern is the total number of its instances in SDB. The set of all instances of a pattern P in SDB is denoted as $\text{Inst}(P)$.

Each instances of a pattern correspond to a remainder of sequence [3]. The remainder of sequence corresponding to an instance is the sub-sequence contained in the sequence which is containing the instance. Assume p is an instance and q is the remainder of sequence corresponding to p . The result of $p++q$ is a suffix of the sequence containing p in SDB. The set of remainder of sequence corresponding to pattern P 's all instances is denoted as $\text{RS}(P)$.

As an example, a sequence database is shown in Table 1. The set $\text{Inst}(A, C)$ is the set of triples $\{(1,2,4), (1,6,8), (2,2,4), (2,7,9), (3,2,4), (3,6,8), \{3,10,12\}, \{4,4,6\}, \{5,2,4\}\}$. The remainder of sequence of instance $(1, 2, 4)$ is $\{A,A,B,C,E,D,F\}$ which is a suffix of $S1$.

Table 1. Sequence database

Identifier	Sequence
$S1$	A,A,B,CA,A,B,C,E,D,F
$S2$	A,A,B,C,B,A,A,B,C
$S3$	$A,A,B,CA,A,B,C,A,A,B,C,E,D,F$
$S4$	B,C,A,A,B,C,E,D,F
$S5$	A,A,B,C,E,C,B,D,F

3.3 Apriori Property and Closed Pattern

Theorem 1. If P is not frequent then its extension are also not frequent.

Definition 5. Consider a pattern P and its super-sequence Q . Instance $I_P(seq_P, start_P, end_P)$ of P corresponds to an instance $I_Q(seq_Q, start_Q, end_Q)$ of Q iff $seq_Q=seq_P$ and $start_P \geq start_Q$ and $end_P \leq end_Q$.

Definition 6. A frequent pattern P is closed if there exists no super-sequence Q s.t.: ① P and Q has the same support ② every instance of P corresponds to a unique instance of Q .

Definition 7. For a pattern P , its set of suffix extension events is defined $\forall s$ the set of length-1 items e where $sup(P ++e)=sup(P)$.

Definition 8. An event e is an infix extension of a pattern P iff \exists a super-sequence Q where: ① $first(P)=first(Q)$ ② \forall event $ev_l \in \text{erasure}(Q, P)$, $ev_l=e$ ③ $sup(P)=sup(Q)$ ④ every instance of P corresponds to a unique instance of Q . ⑤ $\forall rs_P \in \text{ERS}(P)$. $\exists rs_Q \in \text{ERS}(Q)$, $rs_P=rs_Q$.

Definition 9. For a pattern P , its set of prefix extension events is defined as the set of length-1 items e where $sup(e ++P)=sup(P)$.

Theorem 2. If there exist no prefix, infix and suffix extension event $w.r.t$ a pattern P , P must be a closed pattern, otherwise P must be non-closed.

3.4 Pruning Strategies

Suppose there is a pattern list, denoted as S-List, which contains all frequent 2-patterns with size of 2.

Lemma 1. Given a frequent pattern s and its s-extension candidate set CS_n . For each pair of items a, b in CS_n , if $(s->a->b)$ is infrequent, then b must not be a frequent s-extension item pattern $(s->a)$. And for any pattern $(s->s'->a)$, b must not be a frequent s-extension item, too.

Lemma 2. Given a pattern P , if there exists an infix extension event e $w.r.t$ a pattern P and $e \notin rs_P$ ($rs_P \in \text{ERS}(P)$), pattern P can be stopped growing.

4 Algorithm

4.1 Generation Frequent Pattern

In PRISM, the total SDB and all sequences in the SDB are divided into a lot of blocks. PRISM utilizes one positive integer which is denoted as PI to represent all the positional information of mined subsequence's last event in a block. A prime factor of PI represents the positional information of a mined subsequence's last event in a block and belongs to the set which is denoted as G of consecutive prime integer. The minimum element of G is 2. All the elements in G are arranged in increasing order. The i th element and the length of G are respectively denoted as $G[i]$ and $|G|$. Assume that K is the subsequence, which has been processed by PRISM and a is a single element. When PRISM processes sequence extension from K to $K++a$, the greatest common divisor, which is denoted as \gcd , of SEB (sequence blocks) of K and SEB of a is used to obtain the information of the sequence containing both K and a in the SDB. And then POB (position blocks) of K , which is denoted as $\text{POB}(K)$ is converted to obtain the positional information of the event, which is possible the positional information of $K++a$'s last event in all the blocks of each sequence. Finally, the greatest common divisor of converted POB of K and $\text{POB}(a)$ is the $\text{POB}(K++a)$. Assume P is a set of positive integer and is formed as $\{P_1, P_2, \dots, P_n\}$. And Q is a set of positive integer and is formed as $\{Q_1, Q_2, \dots, Q_n\}$. Then $\gcd(P, Q) = \{\gcd(P_1, Q_1), \gcd(P_2, Q_2), \dots, \gcd(P_n, Q_n)\}$. [6]

The position of subsequence isn't the position of the iterative pattern's instance. The SEB of pattern K 's instance is denoted as $\text{ISEB}(K)$. In this paper, the POB of pattern K 's instance, which is denoted as $\text{IPOB}(K)$, is a set of tuple's sets. And each tuple represents the positional information of the instance in a block. In one tuple, the first, the second and the last element respectively represent the positional information of the first event of K 's instance, the last event of K 's instance and the positional information in which all events in the pattern K don't occur. Assume P is a set of tuple's set. The i th element of the j th set in P is denoted as $P_j[i]$. The i th tuple of the j th set in IPOB of pattern K is formed as $(\text{IPOB}_j[i](\text{first}(K)), \text{IPOB}_j[i](\text{last}(K)), \text{NIPOB}_j[i](K))$. If the length of K is 1, $\text{IPOB}_j[i](\text{first}(K)) = \text{IPOB}_j[i](\text{last}(K))$ and it is the value of $\text{NIPOB}_j[i](K)$ that the product of all element in G is divided by $\text{IPOB}_j[i](\text{last}(K))$. Assume K is formed as $K_1K_2\dots K_k$. The set of $\{\text{IPOB}_j[i](\text{last}(K_1)), \text{IPOB}_j[i](\text{last}(K_2)), \dots, \text{IPOB}_j[i](\text{last}(K_k))\}$ is denoted as $\text{IPB}_j[i](K)$. It is the value of $\text{NIPOB}_j[i](K)$ that the product of all element in G is divided by product of all element in $\text{IPB}_j[i](K)$. Assume $G[L_1]$ is a prime factor of the PI which represents the positional information of event L_1 in the g th block and $G[L_2]$ is a prime factor of the PI which represents the positional information of event L_2 in the f th block ($g < f$) in the same sequence. The set of positive integer is denoted as $\text{SET}(G[L_1], G[L_2])$ and used to obtain the positional information from event L_1 to event L_2 in the sequence. The value of the first element of $\text{SET}(G[L_1], G[L_2])$ equals to the product of G 's elements from $G[L_1+1]$ to $G[|G|]$. The value of the last element of $\text{SET}(G[L_1], G[L_2])$ equals to the product of G 's elements from $G[1]$ to $G[L_2-1]$. The value of i th element of $\text{SET}(G[L_1], G[L_2])$ equals to the product of all elements of G . $\{\text{NIPOB}_j[g](e), \text{NIPOB}_j[g+1](e), \dots, \text{NIPOB}_j[f](e)\}$, $\{\text{IPOB}_j[g](\text{first}(e)), \dots, \text{IPOB}_j[f](\text{first}(e))\}$ and $\{\text{IPOB}_j[g](\text{last}(e)) \dots \text{IPOB}_j[f](\text{last}(e))\}$ are respective denoted as $\text{SEN}(e_j(g, f))$, $\text{SEI}(\text{first}(e)_j(g, f))$ and $\text{SEI}(\text{last}(e)_j(g, f))$. The

greatest common divisor can be used to find the positional information relationship between the instance of a pattern and another pattern. For example, the result of $\gcd(\text{IPOB}_j[i](\text{first}(A)), \text{IPOB}_j[i](\text{first}(B))) = \text{IPOB}_j[i](\text{first}(B))$ means that the positional information of the first event of pattern B 's instance is the positional information of first event of part of pattern A 's instances.

Assume that pattern K is formed as $K_1K_2\dots K_k$. So $\text{IPOB}(K)$ and $\text{POB}(K++a)$ are used to obtain the positional information of subsequence $K++a$'s element K_j, K_k and a . Assume $G[N_1]$ is a prime factor of $\text{IPOB}_j[u](\text{first}(K))$ and $G[N_2]$ is a prime factor of $\text{IPOB}_j[b](\text{last}(K))$ and $G[N_3]$ is a prime factor of $\text{POB}_j[i](K++a)$. $G[N_1]$, $G[N_2]$ and $G[N_3]$ respectively represents the positional information of the first event, the last(K) event and the last event of a subsequence $K++a$. $G[N_3]$ is a prime factor of $\text{IPOB}_j[i](\text{last}(K++a))$ if and only if the result of $\gcd(\text{SET}(G[N_2], G[N_3]), \text{SEI}(\text{last}(a)_j(b, i)))$ and $\gcd(\text{SET}(G[N_1], G[N_2]), \text{SEI}(\text{last}(a)_j(u, b)))$ is not the set, in which all the elements are more than 1 and the result of $\gcd(\text{SET}(G[N_2], G[N_3]), \text{SEN}(K_j(b, i)))$ equal to $\text{SET}(G[N_2], G[N_3])$. And then $G[N_1]$ is a prime factor of $\text{IPOB}_j[u](\text{first}(K++a))$. If there do not exist in any instances of pattern $K++a$ in a sequence, the prime factor of an element of $\text{ISEB}(K++a)$, which represent the positional information of $K++(a)$'s instance in the sequence, equal to 1.

Consider the example sequence database shown in Table 1, Let $G = \{2, 3, 5, 7\}$. $\text{POB}(A)$ and $\text{POB}(A \rightarrow B)$ can be obtained by PRISM. $\text{POB}(A) = \{\{6,6,1\}, \{6,15,1\}, \{6,6,6,1\}, \{35,1,1\}, \{6,1,1\}\}$, $\text{POB}(A \rightarrow B) = \{\{5,5,1\}, \{5,14,1\}, \{5,5,5,1\}, \{1,2,1\}, \{5,5,1\}\}$. $\text{IPOB}(A \rightarrow B)$ can be obtained by CIPMPBE. $\text{IPOB}(A \rightarrow B) = \{\{3,5,7\}, \{3,5,7\}, \{1,1,210\}\}, \{\{3,5,7\}, \{5,7,1\}, \{1,1,210\}\}, \{\{3,5,7\}, \{3,5,7\}, \{3,5,7\}, \{1,1,210\}\}, \{\{7,1,6\}, \{1,2,105\}, \{1,1,210\}\}, \{\{3,5,7\}, \{1,1,210\}, \{1,1,210\}\}\}$.

Let $X = \otimes S_X$, with $S_X \subseteq G$. Then $\|X\|_G = |S_X|$ [6]. Note that $\|1\|_G = 0$. For example, $\|21\|_G = \{3, 7\} = 2$. Assume that pattern A of $\text{IPOB}_j[i](\text{last}(e))$ is $\{\{a_1, a_2\} \dots \{a_{(n-1)}, a_n\}\}$. The support of A is $\|a_1\|_G + \|a_2\|_G + \dots + \|a_n\|_G$. So $\text{sup}(A \rightarrow B) = \|5\|_G + \|5\|_G + \|1\|_G + \|5\|_G + \|7\|_G + \|1\|_G + \|5\|_G + \|5\|_G + \|5\|_G + \|1\|_G + \|1\|_G + \|2\|_G + \|3\|_G + \|1\|_G + \|1\|_G = 9$.

4.2 Judgment Closed Pattern

When CIPMPBE obtains a frequent pattern, CIPMPBE must makes a judgment whether the pattern is a closed pattern. So CIPMPBE judge whether the pattern has prefix, infix or suffix extension event. The method is shown as following. Assume that Suf , Inf and Pre are three sets in which every element is a frequent pattern having i evens and satisfy respectively following conditions: ① $\forall a, b \in Suf$, $\text{erasure}(a, \text{last}(a)) = \text{erasure}(b, \text{last}(b))$. ② $\forall a, b \in Inf$, $\text{erasure}(a, a[n]) = \text{erasure}(b, b[n])$ ($1 < n < i$). ③ $\forall a, b \in Pre$, $\text{erasure}(a, \text{first}(a)) = \text{erasure}(b, \text{first}(b))$.

Assume that A_k and A_n are two frequent patterns. Firstly, the conclusion of $\gcd(\text{ISEB}(A_k), \text{ISEB}(A_n))$ equal to $\text{ISEB}(A_k)$. It represents that the sequence containing A_k must contains A_n . Then there are three states as following.

① Assume that A_k and A_n are two elements of Suf . The $\text{last}(A_n)$ is a suffix extension event of A_k if and only if the following condition is satisfied: (1) the result of $\gcd(\text{IPOB}_j[i](\text{first}(A_k)), \text{IPOB}_j[i](\text{first}(A_n)))$ is equivalent to $\text{IPOB}_j[i](\text{first}(A_k))$. (2) The position of $\text{last}(A_n)$ in A_n 's instance is in back of the position of $\text{last}(A_k)$ in A_k 's instance. And the first events of the two instances are in the same position. (3) There is

only one $\text{last}(A_k)$ of which position is between $\text{first}(A_n)$ and $\text{last}(A_n)$. In table 1, an event C is a suffix extension event of pattern AB .

② Assume that A_n and A_k are two elements in Inf and satisfy $\text{erasure}(A_n, A_n[lj]) = \text{erasure}(A_k, A_k[lj])$. There are three states as follow: (1) Assume $\text{gcd}(\text{IPOB}_j[i](\text{first}(A_n)), \text{IPOB}_j[i](\text{first}(A_k)))$ is not equivalent to the value of $\text{IPOB}_j[i](\text{first}(A_k))$. But the result of $\text{gcd}(\text{IPOB}_j[i](\text{last}(A_n)), \text{IPOB}_j[i](\text{last}(A_k)))$ is equivalent to $\text{IPOB}_j[i](\text{last}(A_k))$ and $A_k[lj]=A_n[lj]=A_n[lj]$ ($1 \leq j < i$). The $\text{first}(A_n)$ is an infix extension event of A_k if and only if for every instance of A_k , there is an instance of A_n , which satisfy that the positional information of $\text{first}(A_n)$ is in front of the positional information of $\text{first}(A_k)$. And the last events of the two instances are in the same position. In table 1, an event A is a suffix extension event of pattern AB . (2) Assume that the result of $\text{gcd}(\text{IPOB}_j[i](\text{first}(A_n)), \text{IPOB}_j[i](\text{first}(A_k)))$ equal to $\text{IPOB}_j[i](\text{first}(A_k))$. And the result of $\text{gcd}(\text{IPOB}_j[i](\text{last}(A_k)), \text{IPOB}_j[i](\text{last}(A_n)))$ is equivalent to $\text{IPOB}_j[i](\text{last}(A_k))$. Assume A_k is formed as $B_1 \dots B_{j-1} B D \dots B_i$ and A_n is formed as $B_1 \dots B_{j-1} C D \dots B_i$. Assume $G[n]$ which is a prime factor of $\text{IPOB}_j[N](\text{first}(A_k))$ and $G[m]$ which is a prime factor of $\text{IPOB}_j[M](\text{last}(A_k))$ are respectively the positional information of the first event and the last event of A_k 's an instance. So $\text{gcd}(\text{SEB}(G[n], G[m]), \text{SEI}(\text{last}(B_{j++}B)(N, M)))$ and $\text{gcd}(\text{SEB}(G[n], G[m]), \text{SEI}(\text{last}(B_{j++}C)(N, M)))$ are used to find the order of event D and C in this instance. C is an infix extension event of A_k if and only if the position of B is front of C 's position in all the instances of A_k , or the position of B is in back of C 's position in all the instances of A_k . In table 1, an event D is a suffix extension event of pattern EF . (3) Assume that the conclusion of $\text{gcd}(\text{IPOB}_j[i](\text{first}(A_n)), \text{IPOB}_j[i](\text{first}(A_k)))$ is equivalent to $\text{IPOB}_j[i](\text{first}(A_k))$. And the result of $\text{gcd}(\text{IPOB}_j[i](\text{last}(A_k)), \text{IPOB}_j[i](\text{last}(A_n)))$ is unequal to $\text{IPOB}_j[i](\text{last}(A_k))$. The $\text{last}(A_n)$ is an infix extension of A_k if and only if the position of $\text{last}(A_n)$ of A_n 's instance is in front of the position of $\text{last}(A_k)$ of A_k 's instances. And the first events of the two instances are in the same position. In table 1, an event B is an infix extension event of pattern AC .

③ Assume that A_k and A_n are two elements of Pre . The $\text{first}(A_n)$ is a prefix extension event of A_k if and only if it satisfy the two following condition: (1) The outcome of $\text{gcd}(\text{IPOB}_j[i](\text{last}(A_k)), \text{IPOB}_j[i](\text{last}(A_n)))$ equal to $\text{IPOB}_j[i](\text{last}(A_k))$. (2) For every instance of A_k , there is a instance of A_n which satisfy that there is only one $\text{last}(A_n)$ of which position is in front of the position of $\text{last}(A_k)$. And the last events of the two instances are in the same position. In table 1, an event A is prefix extension event of pattern A .

4.3 Pruning Searching Place

Assume that a pattern A_n has an infix extension event e . And $G[m]$ and $G[n]$ are respectively prime factor of $\text{IPOB}_j[M](\text{last}(A_n))$ and $\text{IPOB}_j[N](\text{first}(A_n))$. $G[m]$ and $G[n]$ represent respectively the positional information of the last event of the i th instance of a pattern A_n and the first event of the $(i+1)$ th instance of a pattern A_n in a sequence. According to Lemma 2, A_n can be stopped growing if and only if $\text{gcd}(\text{SET}(G[m], G[n]), \text{SEI}(\text{last}(e)_j(M, N)))$ is the set in which all the element equal to 1. In table 1, event A is an infix extension event of pattern AB . AB can be stopped growing.

Given a frequent pattern s and its s -extension candidate set CS_n and sa is a not null subset of CS_n . Let $SEP(sa) = \bigcap \{bla \in sa \text{ and } a \rightarrow b \in S\text{-list}\}$, according to LEMMA1, the new s -extension candidate set of sequence $(s \rightarrow sa)$ is $C'S_n = CS_n \cap SEP(sa)$. And, if $(s \rightarrow a \rightarrow b)$ is infrequent according to Lemma 1, delete $(a \rightarrow b)$ from S-list while extend node s and its super-pattern, even if $(a \rightarrow b)$ is frequent. [5]

```

Procedure MineIterativePatterns
Input:
SDB: Sequence Database
min_sup: minimum Support Threshold
Output:
Closed: Closed Iterative Pattern
Methods:
1: Let Csn={ (p,SEB(P),IPOB(p)) | |p|=1 ^ sup(p) ≥ min_sup };
2: Let levelup()=Csn;
3: Let Closed= { };
4: For every a in levelup()
5:   For every b in Csn
6:     a++b;
7:     If(sup(a++b) ≥ min_sup)
8:       add (a++b, ISEB (a++b), IPOB (a++b)) to S-List and leveldown();
9:     End If
10:  End For
11: End For
12: levelup()=leveldown()
13: For every d in levelup()
14:  If preExt(d)={ } ^ sufExt(d)={ } ^ infixExt(d)={ }
15:    add d to Closed;
16:  End If
17:  If ( ∃ i, (i ∈ infixExt(d) ∧ i ∉ RS(d)) )
18:    stop grow d;
19:  End If
20:  Else DSEP (d ,Csn, S-List)
21:    For every e in Csn
22:      d++e;
23:      If(sup(d++e) ≥ min_sup)
24:        add (d++e, ISEB (d++e), IPOB (d++E)) to leveldown();
25:      End If
26:    End For
27:  End Else
28:  levelup()=leveldown();
29: End For

```

Fig. 1. CIPMPBE Algorithm

4.4 Description of Algorithm

Fig.1 shows the pseudo-code of procedure CIPMPBE with all pruning strategies discussed above. In the first step which is shown from line1 to line11 of Fig.1, CIPMPBE constructs Csn of 1-patterns and S-List by sequence extensions from frequent 1-patterns to frequent 2-patterns. In the Second step which is shown in line 14 of Fig.1, CIPMPBE judges whether the obtained frequent pattern has infix extension

event, prefix extension event or suffix extension event. In the third step which is shown from line17 to line19 of Fig.1, Lemma 2 is used to prune search space of the pattern. In line20 of Fig.1 DSEP is used to prune infrequent event. In the final step which is shown from line 21 to line 27 of Fig.1, CIPMPBE generate new frequent pattern which is to return the second step.

5 Experimental Evaluations

To test the performance of CIPMPBE, experiments were made upon a Pentium4 3.0GHz PC with 2GB main memory running Windows XP Professional operation system and Visual studio C++ 6.0. Experiments utilize two synthetic datasets and two real datasets. Synthetic dataset is generated by using the data generator provided by IBM which has the parameters number of sequence (D), average number of events per sequence(C), the number of different events (N), the average items per sequence (T) and the average number of events in the maximal sequence(S). Two really synthetic datasets which are used in experiment are D5C20N10S20 and D5C20T20S20. Gazelle and TCAS datasets are also used in the experiment.

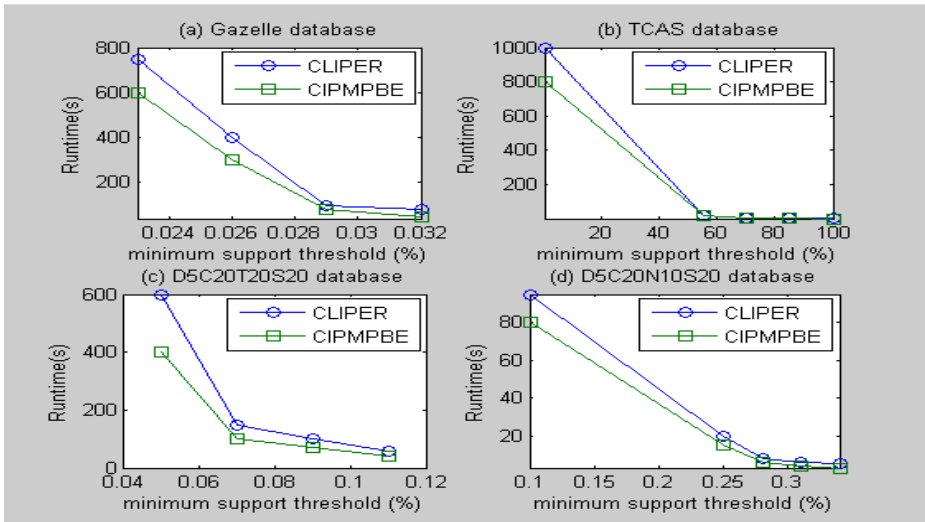


Fig. 2. Performance comparison: varying min_sup

A set of experiments were performed for studying the performance of the CIPMPBE by compare it with CLPER. As shown in Fig.2, the performance is effectively improved in all cases. The CLIPER must be construct a great number of projection databases and scan the projection database repeatedly. So time efficiency of our algorithm is better than CLIPER.

6 Conclusions

In this paper, we have proposed a novel algorithm called CIPMPBE for mining closed iterative pattern. CIPMPBE is based on CLIPER and utilized notion of primal block encoding to process growing of pattern. CIPMPBE makes use of two kinds of pruning strategy to reduce searching space and improve time efficiency.

Closed iterative pattern mining approach opens several research opportunities. In the future, we use CIPMPBE to find the software defects. Improving the efficiency of time and memory is also considered.

References

1. Agrawal, R., Srikant, R.: Mining sequential patterns. In: 11th International Conference on Data Engineering (ICDE 1995), vol. 6, pp. 3–14 (1995)
2. ITU-T. ITU-T Recommendation Z.120: Message Sequence Chart (MSC) (1999)
3. Lo, D., Khoo, S.-C., Liu, C.: Efficient Mining of Iterative Patterns for Software Specification Discovery. In: KDD (2007)
4. Gouda, K., Hassaan, M., Zaki, M.J.: PRISM: An effective approach for frequent sequence mining via prime-block encoding. *Journal of Computer and System Science* (2009)
5. Xu, Y., Zhang, L., Ma, Z., Li, L., Chen, X.: Mining Sequential Pattern Using DF2Ls. In: Fifth International Conference on Fuzzy System and Knowledge Discovery (2008)
6. Mannial, H., Toivonenm, H., Verkamo, A.I.: Discovery of frequent episodes in event sequences. *DMKD 1*, 259–289 (2006)
7. Wang, J., Han, J.: BIDE: Efficient mining of frequent closed sequences. In: ICDE (2004)
8. Ayres, J., Gehrke, J., Yiu, T., Flannick, J.: Sequential Pattern Mining using A Bitmap Representation. In: Proceedings of ACM SIGKDD on Knowledge discovery and data mining, pp. 429–435 (2002)
9. Olender, K., Osterweil, L.: Ceil: A sequencing constraint language for automatic static analysis generation. *IEEE TSE 16*, 268–280 (1990)
10. Pei, J., Han, J., Mortazavi-Asl, B., Pinto, H., Chen, Q., Dayal, U., Hsu, M.-C.: Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth. In: ICDE (2001)
11. Gouda, K., Hassaan, M., Zaki, M.J.: PRISM: A prime-encoding approach for frequent sequence mining. In: IEEE Int'l Conf. on Data Mining (October 2007)
12. Han, J., Pei, J., Mortazavi-Asl, B., Chen, Q., Dayal, U., Hsu, M.-C.: Freespan: Frequent pattern-projected sequential pattern mining. In: KDD (2000)

Pool-Based Active Learning with Query Construction

Shanhong Zhang, Jian Yin, and Weizhao Guo

Department of Computer Science, SUN YAT-SEN University
GuangZhou, China

jt_zsh@163.com, issjyin@mail.sysu.edu.cn, weizhao1758@vip.qq.com

Abstract. Active learning is an important method for solving data scarcity problem in machine learning, and most research work of active learning are pool-based. However, this type of active learning is easily affected by pool size, and makes performance improvement of classifier slow. A novel active learning with constructing queries based pool is proposed. Each iteration the training process first chooses representative instance from pool predefined, then employs climbing algorithm to construct instance to label which best represents the original unlabeled set. It makes each queried instance more representative than any instance in the pool. Compared with the original pool based method and a state-of-the-art active learning with constructing queries directly, the new method makes the prediction error rate of classifier drop more fast, and improves the performance of active learning classifier.

Keywords: active learning, pool-based, construct query, climbing algorithm.

1 Introduction

Active learning is an important method to solve the problem of lack of data labels in machine learning and data mining, which improves performance of classifiers by choosing a small amount of valuable sample to label. Many classification problems in real world such as webpage classification (or voice recognition, image classification) label a large number of samples for supervised learning. Usually these samples are labeled by human experts, which is a costly and time-consuming progress, while massive unlabeled samples are easily to get. Active learning takes the initiative to ask labels of a small quantity of unlabeled sample and reduce the cost of labeling largely.

Pool-based method is one of the hot research topics of active learning and a lot of research methods in this regard have been proposed in recent years [1, 2, 3]. Most previous work of pool-based active learning assume a pool filled with unlabeled data, the most uncertain data is most informative or valuable and select it as representative to label during learning. We use informative and valuable interchangeably following. In real life, it's very likely some valuable data don't exist in the pool, so can't guarantee to get the most valuable data from pool which may results imperfect final classifier. For example, data in a pool has 10 features and each feature value is 0 or 1. So the pool size can be up to 2^{10} with no repeated data. But it's hard to collect 2^{10} data simultaneously in reality, and it's possible some representative data don't exist in the pool. Charles [4] proposed an active learning method with direct query construction, which

constructed the most informative data heuristic on the base of a randomly generated data and effectively avoid the restriction of the pool, yet it's difficult to construct data representing the original pool and quite likely construct outliers or noise. As is shown in figure 1, A is more informative than B while obviously B is more representative of the entire data set, so labeling B enhances performance of active learning classifier more than A. Directly constructed data may also be difficult to label. For example, constructed data is expressed as (35, pregnancy, breast lumps, and male) in binary class setting with breast cancer classification property. It will be difficult for an oracle to label this constructed data.

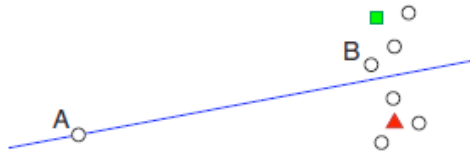


Fig. 1. *Square and triangle* are labeled data of different classes, *rounds* are unlabeled data and *A* is a randomly constructed data which is on the classifying line

In this paper we propose a novel pool-based active learning with query construction CQBP (Construct Queries Based Pool). We select most informative data from pool that best reflect the distribution of unlabeled data in pool by giving a weight to each unlabeled data, and utilize climbing algorithm to construct representative data of original pool based on the selected data to label.

The rest of the paper is organized as follows. Section 2 makes a brief introduction to traditional pool-based active learning. Section 3 details our pool-based active learning method with query construction CQBP. Section 4 compared our method CQBP with traditional pool-base method and ADQ through a large number of experiments, and analyzed effective of CQBP. Section 5 summarized the full-text briefly.

2 Pool-Based Active Learning

In many practical problems, massive unlabeled data are easily to get, which promote pool-based active learning. The basic principle is assuming a small set of labeled data L , a large pool of unlabeled data UL , and generally assumes that the pool is closed. Pool-based active learning first train a classifier on L . Each iterate classifier selects the most informative data from pool based on result of measure, deletes selected data from UL , queries an oracle for label and adds the labeled data to L . Then retrain classifier on L and iteration till meet terminate condition.

Pool-based method has been used to many practical problems because of its simplicity and effectiveness, such as text classification [1], information extract [2], audio recognition and retrieval [3], voice recognition and diagnosis of cancer. Briefly describe below some of the definitions used in this paper. Each data is represented as a d -dimensional feature vector. For i -th data $X_i = (x_{i1}, x_{i2}, \dots, x_{id})$, x_{ij} ($1 \leq j \leq d$) represents value of j -th feature of X_i . Each feature can be numeric, categorical, etc.

The distance between two data is defined in Eq. (1) and the smaller the distance the more similar.

$$dist(i, j) = \frac{\sum_{f=1}^d \delta_{ij}^{(f)} d_{ij}^{(f)}}{\sum_{f=1}^d \delta_{ij}^{(f)}} . \quad (1)$$

In Eq. (1), if x_{if} or x_{jf} is missing, or $x_{if}=x_{jf}=0$ and feature is asymmetric binary variables, $\delta_{ij}^{(f)} = 0$, otherwise $\delta_{ij}^{(f)} = 1$. $d_{ij}^{(f)}$ is contribution of f to distance between i and j , and is calculated below according to its type.

- 1) If f is numeric, $d_{ij}^{(f)} = \frac{|x_{if} - x_{jf}|}{\max_h x_{hf} - \min_h x_{hf}}$, h takes over all data with non-missing f .
- 2) When f is binary or categorical, if $x_{if} = x_{jf}$, then $d_{ij}^{(f)} = 0$, else $d_{ij}^{(f)} = 1$.

3 Pool-Based Active Learning with Query Construction

Pool-based active learning with query construction consists of two steps, which are select step and step of constructing data. We will introduce them below in detail.

3.1 Select Representative Data from Pool

Pool-based active learning selects representative data from pool by a certain way. Proposed choosing strategies come down into six methods, which are uncertainty sampling, query-by-committee, expected model change, variance reduction and fisher information ratio, estimated error reduction and density-weighted methods. Uncertainty sampling[7] is the most widely used selection strategy, it selects most uncertain data from pool and consider it most informative. But this method doesn't consider distribution of data in pool and when the most uncertain data is exactly an outlier, performance of classifier will drop.

V. Chaoji[8] used LOF (Local Outlier Factor) standard to determine whether a data in pool is outlier. Suppose data $x \in D$, given user-defined threshold $Minpts$ neighbor set of x is defined in Eq. (2).

$$N_{Minpts}(x) = \left\{ o \in D \mid dist(x, o) \leq dist(x, x_{Minpts}) \right\} . \quad (2)$$

In Eq. (2), x_{Minpts} is the $Minpts$ -th nearest data to x , so density of x is defined in Eq. (3).

$$density_{Minpts}(x) = \left(\frac{\sum_{o \in N_{Minpts}(x)} dist(x, o)}{|N_{Minpts}(x)|} \right)^{-1} . \quad (3)$$

From Eq. (2) we can see that the smaller the density of data, the more likely it is to outlier. This method requires user to set thresholds. Generally it's difficult to find a suitable threshold, and non-outlier is not necessarily the most informative. When data in pool are dispersive or form a number of clusters, it's hard to select the representative data from pool.

Our CQBP prefers data which both reflects the distribution of data in pool and are most informative. As we focus on query construction in binary classification setting, we set Eq. (4) as a standard to select and construct data.

$$x^* = \underset{x}{\operatorname{arg\,min}} \left| u(x) - 0.5 \right| \times \left(\frac{1}{n} \sum_{u=1}^n dist(x, x^u) \right) . \quad (4)$$

In Eq. (4), x is selected from current pool or constructed. Here it's the representative data of current pool. $u(x)$ is uncertainty value of x , and the smaller $|u(x) - 0.5|$ the more uncertain x that is the more informative in binary classification setting. In multi-classification setting, $|u(x) - 0.5|$ can be replaced by other methods of estimating data

such as information entropy. n is the number of data in pool. $\frac{1}{n} \sum_{u=1}^n dist(x, x^u)$ is

the average distance between x and other data in pool and $dist(x, x^u)$ is calculated

by Eq. (1). The minimum value of $\left| u(x) - 0.5 \right| \times \left(\frac{1}{n} \sum_{u=1}^n dist(x, x^u) \right)$ makes the data

most informative while best reflects distribution of data in pool.

3.2 Construct Representative Data of Original Pool

In traditional pool-based active learning, representative capacity of current pool to original pool diminishes with the downsizing of it, resulting speed of performance improvement slower and slower, especially when the most informative data isn't in original pool. Despite directly constructing data can break the limit of pool size, it's hard to construct representative data of original pool and easily construct outlier or noise. In order to make impact of pool size as small as possible and consider distribution of data in original pool simultaneously, we propose a novel pool-based active learning method with query construction CQBP.

CQBP constructs representative data of original pool based on selected data. Firstly, active learning classifier selects representative data from current pool as initial state, then changes one property value of initial data each time, without changing categories property and this produces a data set that including initial data. Select the most representative data of original pool from the data set using Eq. (4), and then take

the selected data as initial state. Climbing process doesn't stop until the selected data doesn't change any more and the resulting data is the most representative data to label.

3.3 Algorithm Design

The algorithm CQBP proposed in this paper is described in detail in this section. The input of CQBP consists of a whole data set X which is divided into a labeled data set L , an original pool O and a test data set T , an active learning classification model M and termination condition σ . We use U to express the current pool, whose initial value is equal to O . A data x is expressed as (x_1, x_2, \dots, x_d) , consisting of d categories properties. In this paper, we train a classifier on the whole data set to simulate real world oracle. The output of CQBP is the final labeled data set L . The pseudo-code is listed in Figure 2.

```

Input:  labeled data set  $L$ , initial pool  $O$ , current
pool  $U$ , test data set  $T$ , active learning classification
model  $M$  and termination condition  $\sigma$ .
Output: final labeled data set  $L$ .
Begin:
Repeat:
  1) Train  $M$  on  $L$ ;
  2) Select representative data  $x^*$  from  $U$  using
     Eq. (4),  $U=U-x^*$ ;
  3)
     i.   For  $n=1$  to  $d-1$ 
           Begin
                $a_n=x^*$ ;
               Select a possible value of  $x_n$  from  $L$ , and
               assign to  $n$ -th attribute of  $a_n$ ;
           End
     ii.  Select the most representative data of  $O$  from
            $x^*$ ,  $a_1, \dots, a_n$  and assign to  $x$ ;
     iii. If  $x$  is not equal to  $x^*$ ,  $x^*=x$  and go to 3);
  4) Get label of  $x$  from oracle;
  5)  $L=L \cup \langle x, \text{label}(x) \rangle$ ;
Until  $\sigma$  is met.
Output  $L$ .
End.

```

Fig. 2. Pseudo-code of CQBP

M is trained on L first, and each iteration we select the representative data from U and construct most representative data of O by climbing algorithm. The resulting constructed data gets its label from an oracle, removed from U and joined into L . Iteration process stops when termination condition σ is met, and then outputs L .

4 Experiment

4.1 Experiment Setting and Data Sets

In this section, we will prove our CQBP is more effective than traditional pool-based algorithm, which we call Pool for convenience and directly construct data algorithm ADQ[4]. We choose 12 real-world data sets from the UCI Machine Learning Repository[9] to compare our CQBP with Pool and ADQ. These datasets have discrete attributes and binary class without missing values. Information about these data sets is listed in Table 1. In our experiment, each data set is divided into three disjoint subsets, 2 of it is initial training set, 25 percents is testing set, and the remaining is unlabeled set, that is pool in active learning process.

Experiments are conducted on the UCI Machine Learning Repository [9]. An ensemble of 100 j48 decision trees Bagging (implemented in Weka) is used as active learner in this paper. We repeat the experiment 20 times, 50 runs for each data set, and record average prediction error rate of classifiers on test data set.

Table 1. Data sets

Data Set	Attributes	Size	Class Distribution (P/N)
Breast-cancer	9	286	201/58
Breast-w	9	699	458/241
Colic	22	368	232/136
Credit-a	15	690	307/383
Credit-g	20	1000	700/300
Diabetes	8	768	500/268
Heart-statlog	13	270	150/120
Hepatitis	19	155	32/123
Ionosphere	33	351	126/225
Tic-tac-toe	9	958	332/626
Sonar	60	208	97/111
Vote	16	435	267/168

4.2 Experimental Results

Figure 3 displays the results of average prediction error rates of Pool, ADQ and CQBP on 12 data sets from UCI. Horizontal ordinate is number of unlabeled data joined into training set, and vertical ordinate is corresponding error rate of active classifier on testing set. As can be seen from Figure 3, CQBP has the lowest prediction error rate for most data sets.

To quantify the comparison of the learning curve in Figure 3, we list comparison results of t-test with confidence level 95% in Table 2. For each item in Table 2, the algorithm in row is better than algorithm in column on w data sets, equal on t data sets and worse on l data sets. CQBP is better than Pool on 11 data sets, equal on 1 set and worse on 0 sets. This shows that CQBP brings lower predict error rate than traditional pool-based active learning. The same for ADQ, CQBP is better than ADQ on 12 data sets. ADQ is better than Pool only on 6 data sets, equal on 2 sets and even worse on 4

sets. This is due to information of constructed data by ADQ is less than that of data selected from pool, or constructed data does not response distribution of data in pool.

For comparing our CQBP with Pool and ADQ, we record average error rate of 50 iterations of them on each data set and analysis the results statistically, which is displayed in Table 3. Error rate of CQBP reduced 12.65% than Pool on average on 12 data sets and reached 29.11% on Heart-statlog. Error rate of CQBP reduced 12.69% than ADQ on average on 12 data sets and reached 36.24% on Tic-tac-toe. For the same number of queries, active learning classifier trained by CQBP reach higher classification accuracy for joining better unlabeled data.

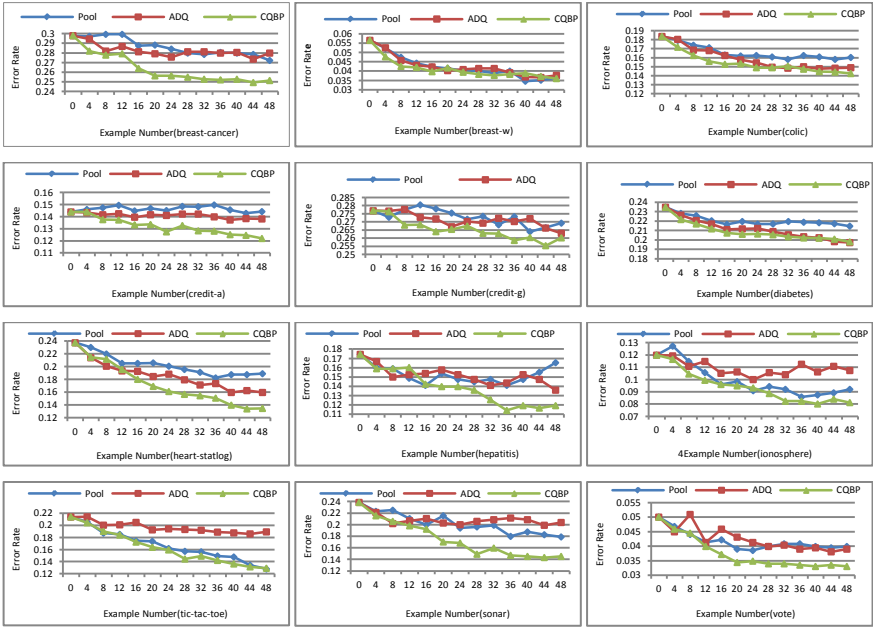


Fig. 3. Error rates of CQBP, Pool and ADQ on 12 data sets. The lower curve, the better.

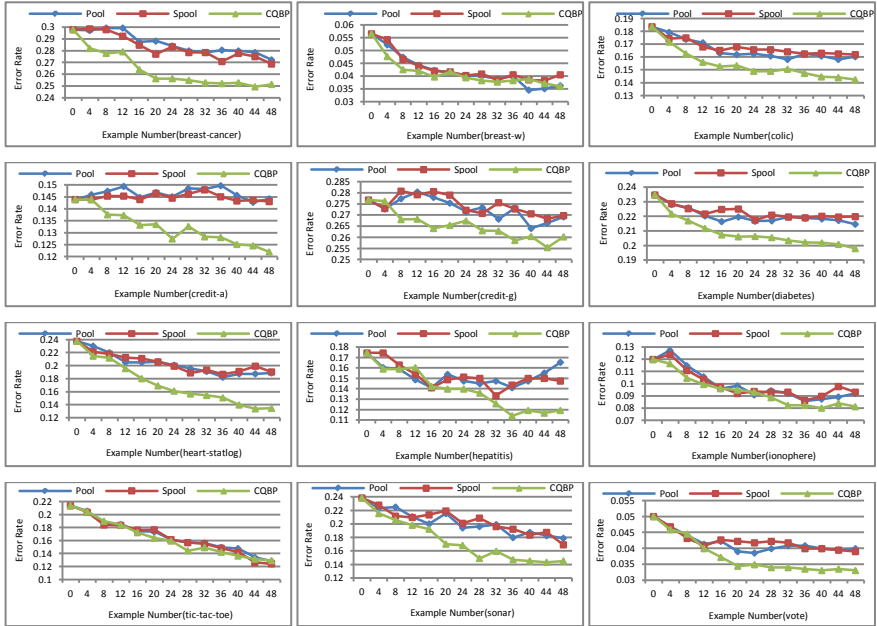
Table 2. Result of t-test on error rate

w/t/l	Pool	ADQ
CQBP	11/1/0	12/0/0
ADQ	6/2/4	

To prove effective of the constructed data by CQBP, we design another algorithm Spool which only selects representative data from pool as CQBP to query without constructing process. Figure 4 displays average prediction error rate of Pool, Spool and CQBP on 12 UCI data sets. Similar to Figure 3, CQBP brings lowest prediction error rate on most data sets, while advantage of Spool is not obvious. This is because Spool still just select data from pool and it doesn't get rid of the limitations of

Table 3. Statistical analysis of the average error rate

Data sets	Error rate% (Pool)	Error rate% (ADQ)	Error rate% (CQBP)	Reduction percentage of error rate% (Pool-CQBP)/Pool	Reduction percentage of error rate% (ADQ-CQBP)/ADQ
Breast-cancer	0.26875	0.275694444	0.251388889	0.064599483	0.088161209
Breast-w	0.036	0.038571429	0.036	0	0.066666667
Colic	0.158152174	0.14673913	0.142391304	0.099656357	0.02962963
Credit-a	0.142485549	0.134104046	0.121965318	0.144016227	0.090517241
Credit-g	0.2706	0.2688	0.259	0.042867701	0.036458333
Diabetes	0.215364583	0.196875	0.196614583	0.087061669	0.001322751
Heart-statlog	0.191911765	0.154411765	0.136029412	0.291187739	0.119047619
Hepatitis	0.165384615	0.138461538	0.123076923	0.255813953	0.111111111
Ionosphere	0.092613636	0.107954545	0.082954545	0.104294479	0.231578947
Tic-tac-toe	0.131666667	0.190833333	0.121666667	0.075949367	0.362445415
Sonar	0.177884615	0.193269231	0.144230769	0.189189189	0.253731343
Vote	0.039449541	0.038073394	0.033027523	0.162790698	0.13253012
Average error rate	0.157521929	0.156982321	0.137362161	0.126452239	0.126933366

**Fig. 4.** Error rates of CQBP, Pool and Spool on 12 data sets. The lower curve, the better

pool-based method. So we construct the most representative data of original pool for further after selecting data from current pool.

5 Conclusions

We introduced a pool-based active learning method with query construction and proposed CQBP algorithm. We propose a novel standard to select and construct data. Large numbers of experiments show that our algorithm makes query in each time more valuable than any data in current pool and break the limitation of the size of pool. We also take into account distribution of data in current pool and avoid constructing outlier or noise, thereby improve performance of active learning classifiers.

Acknowledgments. This research is supported in part by the National Natural Science Foundation of China (NSFC) 61033010, in part by the National Science and Technology Projects of China (2008ZX10005-013), and in part by Science and Technology Projects of Guangdong (2009A080207005, 2009B090300450, 2010A040303004).

References

1. Hoi, S.C.H., Jin, R., Lyu, M.R.: Large-scale text categorization by batch mode active learning. In: Proceedings of the International Conference on the World Wide Web, pp. 633–642. ACM Press (2006)
2. Settles, B., Craven, M.: An analysis of active learning strategies for sequence labeling tasks. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1069–1078. ACL Press (2008)
3. Hauptmann, A., Lin, W., Yan, R., Yang, J., Chen, M.Y.: Extreme video retrieval: joint maximization of human and computer performance. In: Proceedings of the ACM Workshop on Multimedia Image Retrieval, pp. 385–394. ACM Press (2006)
4. Ling, C.X., Du, J.: Active Learning with Direct Query Construction. In: KDD, pp. 480–487 (2008)
5. Lewis, D., Gale, W.: A sequential algorithm for training text classifiers. In: Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 3–12. ACM/Springer (1994)
6. Settles, B.: Active learning literature survey. Technical Report 1648, University of Wisconsin–Madison (2010)
7. Lewis, D.D., Catlett, J.: Heterogeneous Uncertainty Sampling for Supervised Learning. In: Proceedings of the International Conference on Machine Learning, pp. 148–156 (1994)
8. Chaoji, V., Hasan, M.A., Salem, S., Zaki, M.J.: SPARCL: Efficient and Effective Shape-based Clustering. In: ICDM, pp. 93–102 (2008)
9. Asuncion, A., Newman, D.J.: UCI machine learning repository (2007), <http://www.ics.uci.edu/~mllearn/mlrepository.html>

10. Witten, I.H., Frank, E.: *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edn. Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann (June 2005)
11. Du, J., Ling, C.X.: Asking Generalized Queries to Domain Experts to Improve Learning. In: *TKDE* (2010)
12. Baum, E.B., Lang, K.: Query learning can work poorly when a human oracle is used. In: *IEEE International Joint Conference on Neural Networks* (1992)
13. Nguyen, H.T., Smeulders, A.: Active Learning Using Pre-clustering. In: *ICML*, pp. 623–630 (2004)

The Research of Soft Measurement Method Based on Sintering Process Permeability Index

Jinyu Teng and Xiaoxin Zhang

Department of Electronics Information Engineering
Shen Yang Aerospace University, Shen Yang, China
erlian0308@sina.com

Abstract. This paper studies the modeling method of soft measurement based on the permeability of index sintering process. Soft sensor modeling can use three kinds of methods, neural network, fuzzy control and adaptive fuzzy neural network of soft measurement based on subtraction clustering. Through analysis various soft sensor modeling methods and experimental data, Subtraction clustering adaptive fuzzy neural network method has a very good convergence, and also it has highly precision of prediction, smaller test error, so the method very suitable for the engineering application.

Keywords: soft measurement, sintering process, fuzzy control, subtraction clustering.

1 Introduction

The basic idea of the soft measurement model is based on certain optimal criteria to choose a group leading variables; it has close relationship with the variables and it also easy to be measured. By constructing a mathematical relationship to estimate the dominant variables, process object input/output relationship shown in figure 1. y is leading variable, which is difficult to be measured, d_1 is measurable interference variable, d_2 is cannot be measured interference variable, u is measurable control variable, φ is measurable controlled variable. Leading variable y which estimated value can be shown in the equation $y=f(x)$, x is measurable auxiliary variable, $x \subseteq (d_1, u, \varphi)$.

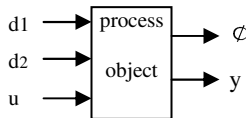


Fig. 1. Process object input/output relationship

Permeability state of sintering raw material layer is a very important state parameter to sintering process, which reflects the degree of difficulty when gas flowing through the sintering raw material layer, and the gas fluctuation and the change rule in

sintering raw material layer, as the same time, permeability state influence mass transfer, heat transfer and physical and chemical reaction of sintering process [1]. Then sintering material combustion is not full, that easy to produce raw layer, if permeability is bad. On the one hand that can decrease sintering block output, on the other hand make sintering blocks contains high percentage of sulfur, will eventually lead to convulsions furnace smelting tumor blocky.

Though analyze the influencing factors of sintering raw material layer permeability that the main factors includes air volume, trolley speed ,ignition temperature, return ore and water quantity. The main factors which affect the permeability as auxiliary variables (The input of the soft measurement model), permeability index as the output of soft measurement model, establish the permeability index soft measurement model.

2 The Soft Measurement Based on Neural Network Modeling

In recent years, the fastest growing and most wide application of a kind of soft measurement technique is based on artificial neural network of soft measurement. Because the system input parameter and output parameter continuous change, so the neural network's input value and the output value should be continuously. In training requirements system output and target output errors small enough that means it needs supervision and training. So, the system of neural network models at least should meet the following three conditions: continuous input, continuous output and supervision training, BP network can meet these requirements. Meanwhile, as one of the most mature network model, it is also the most widely used in practice, therefore, selected layers of BP network [2].

The BP neural network with five input variables and one output variables, hidden layer number of neurons is 26. Due to the output requirements in [- 1 1], so the transmission function of first layer using tansing. The second requirement output linear, so transmission function using purelin. Training function select trainlm because the algorithm has very fast, the training speed for 0.5, learning objectives adopted for 0.001. Data need normalized before neural network training, input/output requirements in [- 1 1]. The training samples have 1104, test samples have 160 and training steps are 10000.

Inspection samples with 160 groups to verify the network performance, the results shown in table 1. Table only sixteen of the group to explain. From soft measurement results can be seen data measurement error is less than 3.03%, system error meet the requirements.

Table 1. Neural network test.

number	actual value	test value	error	number	actual value	Test value	error	number	actual value	test value	error
1	6.5%	6.37%	-2.00%	5	7.3%	7.2%	-1.37%	9	6.6%	6.4%	-3.03%
2	7.7%	7.85%	1.95%	6	8.2%	8.28%	0.98%	10	7.2%	7.39%	2.64%
3	8.4%	8.55%	1.79%	7	7.7%	7.89%	2.47%	11	6.7%	6.87%	2.54%
4	6.5%	6.59%	1.38%	8	6.8%	6.98%	2.65%	12	8.2%	8.38%	2.20%

3 Soft Sensor Modeling Based on Fuzzy Reasoning

Fuzzy logic thinking characteristics are imitating the human brain, which is an effective way to dealing with complex system, also got used in process soft measurement. Based on fuzzy mathematics model of the soft measurement is a kind of knowledge model. This kind of soft measurement method is especially suitable for complex industry process that the tested object presents uncertainty or it is difficult to use conventional mathematics quantitative description. Permeability index measuring nonlinear and uncertainty of material layer in sintering process. There is not ready-made formula can be used, but there are many empirical knowledge can be used, so it can be modeled by using fuzzy inference [3].

Fuzzy reasoning forecasting model and neural network model with the same input and output variables, Input variables elected air volume, the ignition temperature, trolley speed, return ore and water quantity total five variables, permeability index as the output variables.

Using the same inspection data test fuzzy reasoning model. Choose the 16 data from 160 data as comparing show in table 2.

Table 2. Fuzzy model test

number	actual value	test value	error	number	actual value	test value	error	number	actual value	test value	error
1	6.5%	6.79 %	4.46%	5	7.3%	6.73%	-7.8%	9	6.6%	7.48%	13.33%
2	7.7%	7.78%	-1.19%	6	8.2%	7.51%	-8.41%	10	7.2%	7.51%	4.31%
3	8.4%	7.27%	-13.45%	7	7.7%	7.51%	-2.47%	11	6.7%	7.52%	12.24%
4	6.5%	6.74%	3.69%	8	6.8%	7.51%	10.44%	12	8.2%	6.77%	-17.44%

Judging from the results error is bigger, the maximum error is 17.44%, the minimum error is 1.19%, and average error is 8.32%, from that this system error is bigger. If we want to let the error to smaller, we must division interval of input variables of membership functions more detailed. If each input membership functions are divided into 5, then fuzzy rule number will increase to 625, established in 625 rules of fuzzy system calculation at a slower speed, calculation results still cannot reach the accuracy.

4 A Soft Sensor Modeling Based on Fuzzy Neural Network

Using the fuzzy neural network realization soft sensor modeling is to explore new ideas in recent years. On one hand, we can solve some problem of the soft measurement, On the other hand also found a new application field for fuzzy neural network. We can think that neural network to simulate the human brain hardware, and fuzzy technology simulation of the human brain software, the two methods combine to form the fuzzy neural network technology. Fuzzy neural network technology also it can be applied to sintering process of material layer in the modeling of permeability index. This paper introduces fuzzy neural network based on the subtraction clustering type of adaptive neural fuzzy reasoning system. Based on the data modeling method is the biggest characteristic of it [4]. Fuzzy membership functions and the fuzzy rules of the system are not based on experience but it is known by learning. Numerical clustering

analysis is the base of classification and system modeling. The purpose of clustering is to extract features from a lot of data, that to simply show the system behavior. Subtraction clustering is a fast single algorithm used to estimate of a set of data the cluster number and clustering center position. Subtraction clustering algorithm estimates can be used for the beginning of those based on the fuzzy clustering and model identification method optimization repeat process [5].

If subtraction clustering method and ANFIS method combining (SCANFIS), the first is to use the subtraction clustering cluster analysis practical problems get the prototype of system clustering center, then by these clustering center structure a group initial fuzzy model with a order Sugeno (structure identification), and then use the ANFIS good parameter identification method to adjust before parameters and consequent parameters of the rules (parameter identification), this model called SCANFIS model. SCANFIS model input/output parameters, training data and inspection data as same as subtraction clustering method and ANFIS method. In order to select a suitable clustering radius r_a , we took different r_a value respectively clustered, the experimental conclusion show that this model structure is simple and there are 14 rules only. Used SCANFIS methods test 160 group inspection data, and calculated every test sample error; the sixteen of the group is shown in table 3. One of the biggest measurement error was 2.58%, the minimum error is 0.12%, be able to comply with system requirements.

Table 3. Fuzzy neural network model test

number	actual value	test value	error	number	actual value	test value	error	number	actual value	test value	error
1	6.5%	6.48%	-0.31%	5	7.3%	7.31%	0.14%	9	6.6%	6.43%	-2.58%
2	7.7%	7.71 %	0.23%	6	8.2%	8.21%	0.12%	10	7.2%	7.23 %	0.42%
3	8.4%	8.41 %	0.12%	7	7.7%	7.69%	-0.13%	11	6.7%	6.53%	-2.37%
4	6.5%	6.35 %	-2.31%	8	6.8%	6.82 %	0.29%	12	8.2%	8.21%	0.12%

5 Three Soft Measurement Method Comparison

The comparison fuzzy neural network model and the fuzzy model: It input variables of membership functions more careful than the fuzzy model, In the experiment the fuzzy nerve network rule number is 14, but fuzzy model rule number is 81, And the fuzzy neural network model error is smaller. It is difficult to use fuzzy model establish rules to achieve the same error level, even if the rules established but it operation speed are also very slowly. Judging from the test results, the fuzzy neural network model errors achieve $8.4267e-006$, the fuzzy model can only achieve minimum error around 0.01.

The comparison fuzzy neural network model and neural network model, Fuzzy neural network model hidden layer node number is 14, Neural network model hidden layer node number is 26, Fuzzy neural network training a few steps can achieve error requirements, neural network training 5980 steps to the error is 0.001. Because neural network training more steps so the training time is very long, and fuzzy neural network has better generalization than neural network. Table 5.1 is three kinds of soft measurement methods measurement error contrast table, from the analysis of the

testing result, fuzzy neural network model error is minimum and the fuzzy model error is the largest. The running speed of fuzzy neural network model is the fastest, neural network is the slowest.

Table 4. Three soft measurement method comparison

numbe	actual value	neural network test value	fuzzy test value	fuzzy neural network test value	neural net- work error	fuzzy error	Fuzzy neural net- work error
1	6.5%	6.37%	6.79%	6.48%	-2.00%	4.46%	-0.31%
2	7.7%	7.85%	6.78%	7.71%	1.95%	-1.19%	0.23%
3	8.4%	8.55%	7.27%	8.41%	1.79%	-13.45%	0.12%
4	6.5%	6.59%	6.74%	6.35%	1.38%	3.69%	-2.31%
5	7.3%	7.2%	6.73%	7.31%	-1.37%	-7.8%	0.14%
6	8.2%	8.28%	7.51%	8.21%	0.98%	-8.41%	0.12%
7	7.7%	7.89%	7.51%	7.69%	2.47%	-2.47%	-0.13%
8	6.8%	6.98%	7.51%	6.82%	2.65%	10.44%	0.29%
9	6.6%	6.4%	7.48%	6.43%	-3.03%	13.33%	-2.58%
10	7.2%	7.39%	7.51%	7.23%	2.64%	4.31%	0.42%
11	6.7%	6.87%	7.52%	6.53%	2.54%	12.24%	-2.37%
12	8.2%	8.38%	6.77%	8.21%	2.20%	-17.44%	0.12%

From simulation experiment results, Fuzzy neural network model method is better, fuzzy neural network is combination of fuzzy logic and the neural network. Fuzzy neural network makes neural network the "black box" problem to transparent; it is meaning that can use the fuzzy neural network's input/output realize complex rules. Fuzzy neural network learning through plenty of examples of samples, constantly adjust the width and center of the network membership functions, Thus to grasp the information that use the equation is difficult to express in the sample. Fuzzy neural network through the membership function records the learned sample rule, which can master system's complicated nonlinear relation between input and output, so it can made precision and stability of evaluation results the better. Comparison of the other two methods, fuzzy neural network modeling, computation or reasoning is very easy for complex nonlinear object which is difficult to accurately describe.

References

1. Cheng, J.J.: Neural network theory, vol. 5(3), pp. 38–41. Xian university of electronic science and technology press, Xian, China (1995) (in Chinese)
2. Marcos, J., Arauzo, B.: Automatization of a penicillin production process with soft sensors and an adaptive controller based fuzzy systems. *Control Engineering Practice* 12(9), 1073–1090 (2004)
3. Li, H.Q.: The soft measurement technology principle and application, pp. 73–74. Chemical industry press, China (2000) (in Chinese)
4. Wang, S.T.: Fuzzy system, a fuzzy neural network and application design, vol. 4(21), pp. 82–85. Shanghai science and technology literature press, Shanghai, China (2008) (in Chinese)
5. Wang, X.D.: Based on neural network of universal soft measurement technique. *Acta Automatica, China* 24(5), 702–703 (2008) (in Chinese)

A Neighborhood Preserving Based Semi-supervised Dimensionality Reduction Method for Cancer Classification

Xianfa Cai^{1,2}, Jia Wei², Guihua Wen², and Jie Li¹

¹ School of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou 510006, China

² School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China
cxianfa@126.com

Abstract. Cancer classification of gene expression data helps determine appropriate treatment and the prognosis. Accurate prediction to the type or size of tumors relies on adopting efficient classification models such that patients can be provided with better treatment to therapy. In order to gain better classification, in this study, a linear relevant feature dimensionality reduction method termed the neighborhood preserving based semi-supervised dimensionality reduction (NPSSDR) is applied. Different from traditional supervised or unsupervised methods, NPSSDR makes full use of side information, which not only preserves the must-link and cannot-link constraints but also can preserve the local structure of the input data in the low dimensional embedding subspace. Experimental results using public gene expression data show the superior performance of the method.

Keywords: Cancer Classification, Semi-supervised Learning, Side-information, Local Preserving.

1 Introduction

The advent of DNA microarray technique has made it possible to simultaneously measure the expression levels of thousands of genes in a single experiment. Gene expression profiling by microarray technology has been successfully applied to classification and diagnostic prediction of cancers, and accurate prediction of different tumor types has great value in providing better treatment and toxicity minimization on the patients. However one of the challenging tasks is how to identify salient expression genes from thousands of genes in microarray data that can directly contribute to the symptom of disease.

Due to high dimensionality and a small number of noisy samples, gene expression data poses great challenges to the existing machine learning methods and usually results in the known problem of “curse of dimensionality” [1]. During past years, it has drawn a great deal of attention from both biological and engineering fields. Different classification methods have been applied to cancer classification, such as, support vector machines(SVM) [2,3], neural net-works [4], logistic regression [5] and

GA/KNN [6]. More recently, with the development of machine learning, feature extraction has been extensively used in cancer classification, which aims to extract the most representative features with low dimensions from the original data. The most common methods involved in gene data analysis are principal component analysis(PCA), partial least squares(PLS)[7] and independent component analysis(ICA)[8],etc. A systematic benchmarking of these methods is reported in the literature[9]. Besides, lots of nonlinear feature extraction methods have been developed for cancer classification, such as kernel principal component analysis, LLDE[10,11,12],etc.

However in real-world applications, it is usually easy to get unlabeled samples, but difficult to obtain labeled ones because labeling is an expensive and time-consuming job. It leads to the potential waste of valuable classification information buried in unlabeled samples. At the same time, we find that pairwise constraints information also called side-information is more general than label information, since we can obtain side information from label information but not vice verse[13]. So learning with side information is becoming an important area in the machine learning .With this in mind, ' semi-supervised learning' method learning with both labeled and unlabeled data, has attracted much attention in recent years[14]. Zhang et al. [15]proposed semi-supervised dimensionality reduction (SSDR), which can preserve the intrinsic structure of the must-link and cannot-link constraints as well as the unlabeled data in the projected low dimensional space. It can only preserve the global covariance structure but fail to preserve local structure of data. Cevikalp et al. [16]proposed constrained locality preserving projections(CLPP) recently, which can use the must-link and cannot-link information and also use the unlabelled data by preserving local structure. However, there's no reliable approach to determine an optimal parameter t to construct the adjacency graph. So, Wei et al. [17] proposed a method termed the neighborhood preserving based semi-supervised dimensionality reduction algorithm(NPSSDR),which makes full use of side information, not only preserves the must-link and cannot-link constraints but also can preserve the local structure of the input data in the low dimensional embedding subspace. In this paper, to solve the cancer classification problem, we put NPSSDR into the application of gene expression data for cancer classification and experimental results demonstrate the effectiveness on several gene datasets.

The rest of the paper is organized as follows. In section 2, we review the LLE. In section 3, we introduce the proposed NPSSDR method. The experimental results are presented in section 4. In section 5, we conclude our methods and provide some suggestions for future work.

2 Related Work

2.1 Locally Linear Embedding

Locally Linear Embedding [18], proposed by Saul and Roweis, tries to find a nonlinear manifold by stitching together small linear neighborhoods. The main principle of LLE is to preserve local neighborhood relation in both the intrinsic space and embedding one by finding a set of weights that perform local linear interpolations that closely

approximate the data. Each sample in the observation space is a linearly weighted average of its neighbors. In general, the procedures of LLE can be stated in three steps:

Step 1. define neighbors for each data point. In this Letter, we use k-nearest neighbor to model the relationship between nearby points. Euclidean distance is used to define the neighborhood $d_{ij} = \|x_i - x_j\|$.

Step 2. find weights that allow neighbors to interpolate original data accurately.

$$\min \varepsilon (W) = \sum_{i=1}^N \left\| x_i - \sum_{j=1}^N W_{ij} x_j \right\|^2 \tag{1}$$

Step 3. find new data points that minimize interpolation error in lower dimensional space. Compute the low-dimensional embedding Y of X that best preserves the local geometry represented by the reconstruction weights.

$$\min \varepsilon (W) = \sum_{i=1}^N \left\| y_i - \sum_{j=1}^N W_{ij} y_j \right\|^2 \tag{2}$$

For more details of LLE, please refer to [18].

3 NPSSDR

Here we define the side-information based semi-supervised linear dimensionality reduction problem as follows. Suppose we have a set of D -dimensional points $X = \{x_1, x_2, \dots, x_n\}^T, x_i \in R^D$ together with some pairwise must-link(M) and cannot-link(C) constraints as side-information: $M : (x_i, x_j) \in M$ if x_i and x_j belong to the same class; $C : (x_i, x_j) \in C$ if x_i and x_j belong to the different class. Then the next step is to find a transformation matrix A such that the transformed low dimensional projections $Y (y_i = A^T x_i)$ can preserve the structure of the original dataset as well as the side information M and C . That is, points in M should end up close to each other while points in C should end up as far as possible from each other.

Since it's easy to extend to high dimensions, for the convenience of discussion, the one dimensional case is considered here. As for must-link constraints M and cannot-link constraints C , Q_m denotes the intra-class compactness and Q_c denotes the inter-class separability.

Defined to describe the compactness of labeled samples that belong to the same classes, Q_m is measured as follows:

$$\left\{ \begin{aligned} Q_m &= \sum_{(x_i, x_j) \in M \text{ or } (x_j, x_i) \in M} (w^T x_i - w^T x_j)^2 = 2 \sum_i (w^T x_i D_{ii}^m x_i^T w) - 2 \sum_{ij} (w^T x_i S_{ij}^m x_j^T w) \\ &= 2w^T X (D^m - S^m) X^T w = 2w^T X L^m X^T w \\ S_{ij}^m &= \begin{cases} 1, & \text{if } (x_i, x_j) \in M \text{ or } (x_j, x_i) \in M \\ 0, & \text{else} \end{cases} \end{aligned} \right.$$

Where D^m is a diagonal matrix the entries of which are column(or row, since S^m is symmetric) sums of S^m , $D_{ii}^m = \sum_j S_{ij}^m$, $L^m = D^m - S^m$ is a graph Laplacian matrix that is positive semi-definite[19].

Similarly, the inter-class separability for samples of different classes can be termed as follows:

$$\left\{ \begin{array}{l} Q_c = \sum_{\substack{ij \\ (x_i, x_j) \in C \text{ or } (x_j, x_i) \in C}} (w^T x_i - w^T x_j)^2 = 2w^T X(D^c - S^c)X^T w = 2w^T XL^c X^T w \\ S_{ij}^c = \begin{cases} 1, & \text{if } (x_i, x_j) \in C \text{ or } (x_j, x_i) \in C \\ 0, & \text{else} \end{cases} \end{array} \right.$$

Where D^c is a diagonal matrix, $D_{ii}^c = \sum_j S_{ij}^c$, $L^c = D^c - S^c$.

Inspired by the Fisher Linear discriminant analysis criterion[20], we can define a raw objective function in terms of Q_c and Q_m as follows:

$$w^* = \arg \max_w \frac{Q_c}{Q_m} = \arg \max_w \frac{w^T XL^c X^T w}{w^T XL^m X^T w} \quad (3)$$

Eq.(3) achieves the goal of placing samples of the same classes close to each other while samples of different classes as far as possible from one another in the subspace. However, Eq.(3) suffers from singular problem, since the number of labeled samples is often smaller than the dimensionality of samples.

Besides, to make use of a large amount of unlabeled data, we introduce a regularisation term to preserve the neighborhood relationships of input space. We assume that all the neighborhoods of input space are linear as Roweis and Saul[18] did, so each data point can be optimally reconstructed using a linear combination of its neighbors. Hence, our objective is to minimize the reconstruction error according to the Eq.(1):

$$\min \mathcal{E}_1(W) = \sum_{i=1}^N \left\| x_i - \sum_{j, x_j \in N(x_i)} W_{ij} x_j \right\|^2$$

Where $N(x_i)$ represents the neighborhood of x_i .

In the low dimensional space, we want to minimize the cost function as follows according to the Eq.(2):

$$\min \mathcal{E}_2(Y) = \sum_i \left\| y_i - \sum_{j, x_j \in N(x_i)} W_{ij} y_j \right\|^2 = \text{trace}(YMY^T) = \text{trace}(W^T XMX^T W) \quad (4)$$

Where $M = (I - W)^T (I - W)$, I is the identity matrix.

Based on the above analysis, the objective function of NPSSDR is:

$$A_{opt} = \arg \max_A \frac{Q_c}{Q_m + \alpha \mathcal{E}_2} = \arg \max_A \frac{A^T XL^c X^T A}{A^T X(L^m + \alpha M)X^T A} \quad (5)$$

Where scaling parameter α is set to balance the contribution of \mathcal{E}_2 .

To solve the above optimization problem, we use the Lagrangian multiplier and finally we have:

$$XL^c X^T A = \lambda X(L^m + \alpha M)X^T A \tag{6}$$

It's easy to know that the d eigenvectors corresponding to the d largest eigenvalues can form the transformation matrix A .

4 Experimental Results

In this section, we conduct several experiments on five gene expression datasets(described in Table 1) to compare the effectiveness of NPSSDR with other methods. The comparative methods are Baseline(which is simply the nearest neighbor classifier on the original input space) 、PCA 、CLPP and SSDR. The pairwise constraints in the experiments are obtained by randomly selecting pairs of instances from the training set and creating must-link or cannot-link constraints depending on whether the underlying classes of the two instances are the same or not. The nearest neighbor classifier is used to the class testing set after dimensionality reduction. To reduce noise, we first project the datasets to a PCA subspace with the 98% principal components preserved. If without extra explanation, the parameters in NPSSDR are always set to $k=5$ and $\alpha=0.1$. Some notations are described in Table 2.

Table 1. General information of the datasets

Dataset	No. of samples	No. of features	Classes
Colon	62	2000	2
Ovarian	253	15154	2
DLBCL	77	7129	2
Leukemia	72	7129	2
Prostate	136	12600	2

Table 2. Notations used in the experiments

Notation	Meaning
Tr	number of training samples
Te	number of testing samples
d	target dimensionality
k	neighborhood numbers
NOC	number of constraints
α	scaling parameter

4.1 Results on Different Target Dimensionalities

In order to illustrate the effectiveness of NPSSDR on different target dimensionalities, we conduct five experiments. In these experiments, we vary d while set NOC as 100. The results are listed in Fig. 1. All the results are the 10 independent runs to

balance the random effects. In most cases NPSSDR performs best, especially in Fig. 1((a),(b),(c)), followed with SSSDR. Generally speaking, SSSDR achieves the most stable effect, even though NPSSDR achieves the best performance in most of datasets. As the target dimensionality increases, our method's accuracies increase and reach

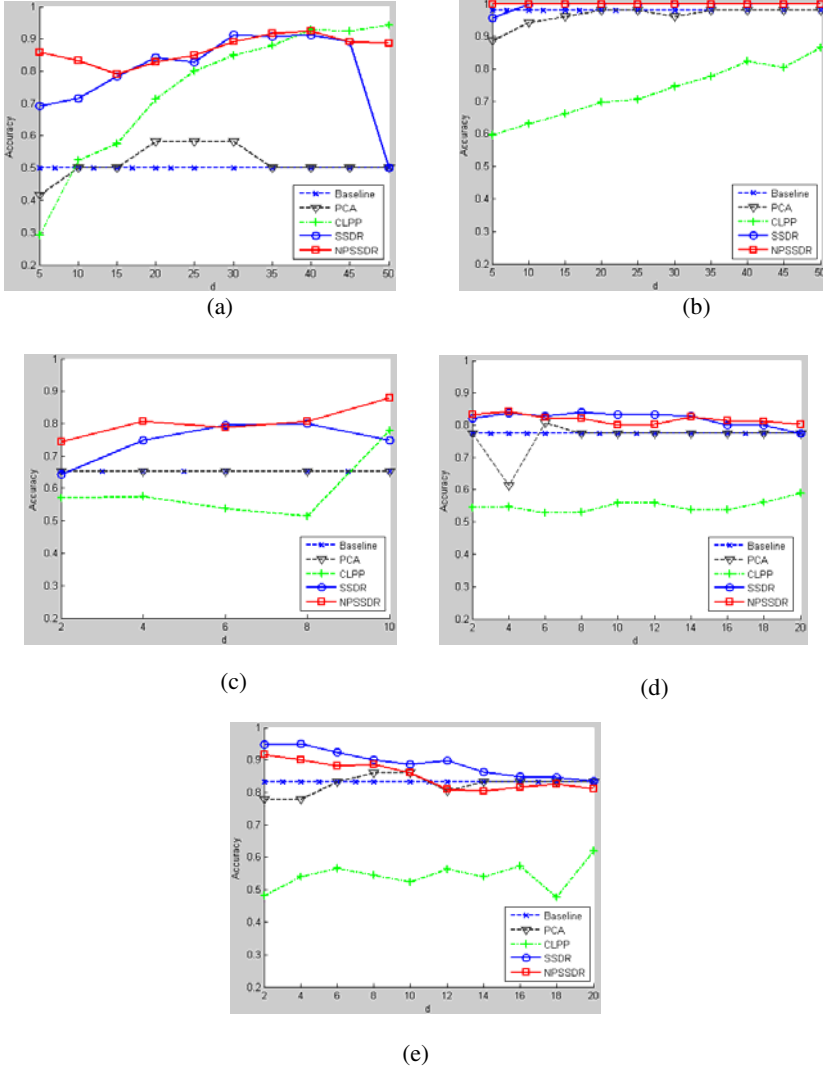


Fig. 1. Classification accuracies of different methods on the different target dimensionalities. (a) DLBCL (Tr=60, Te=17, NOC=100) (b) OVARIAN (Tr=200, Te=53, NOC=100) (c) PROSTATE (Tr=100, Te=36, NOC=100) (d) COLON (Tr=31, Te=31, NOC=100) (e) LEUKEMIA (Tr=36, Te=36, NOC=100)

comparatively high accuracies, while the accuracies of the PCA and CLPP are always in relative low accuracies. Dimensionality reduction can promote the learning results are also demonstrated in Fig.1, for NPSSDR and SSSDR all have higher accuracies than Baseline does. Besides, NPSSDR and SSSDR take PCA as a preprocessing step, but they gain higher precisions than single PCA does, which also demonstrates the effectiveness of these methods.

4.2 Results on Different Number of Constraints

To explore the effect of *NOC* on accuracy, we conduct another five experiments in which *d* is set. To study the effect of NPSSDR compared to other methods, experimental results are listed in Fig.2. Experimental results indicate that NPSSDR can make use of constraints efficiently. As *NOC* increases, NPSSDR and SSSDR gain higher and higher accuracies, on the contrary, CLPP is always in relative low accuracies. In reality, CLPP benefits little from pairwise constraints even when the number of pairwise constraints is relative large. The results demonstrate NPSSDR can make use of pairwise constraints more efficiently than other comparing methods.

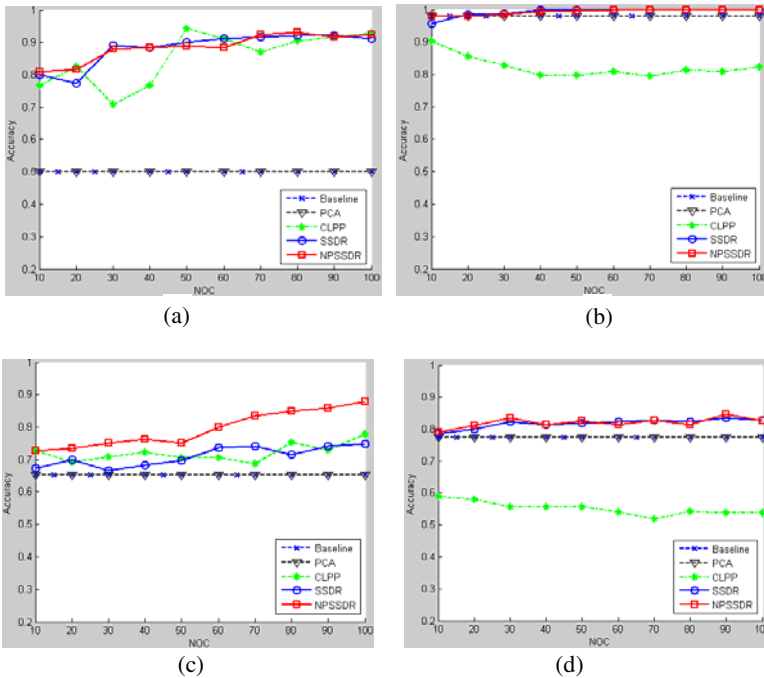
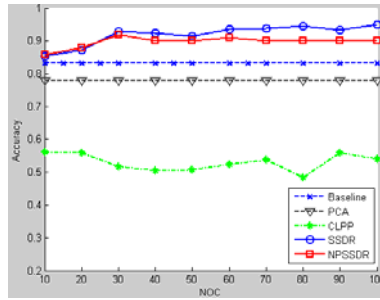


Fig. 2. Classification accuracies of different methods on the different NOC (a) DLBCL(d=40) (b) OVARIAN(d=40) (c) PROSTATE(d=10) (d) COLON(d=14) (e) LEUKEMIA(d=4)



(e)

Fig. 2. (continued)

5 Conclusions and Future Work

In this paper, we conduct cancer classification based on NPSSDR method and obtain good results which demonstrate semi-supervised learning has great value in cancer classification. In comparison with SSSR and CLPP, NPSSDR exploits side information efficiently. Besides, information buried in unconstrained data is sufficiently considered. Thus NPSSDR has always comparatively high classification accuracies with different NOC and target dimensionalities. However, NPSSDR still has some issues for future work. Parameters k and α in experiments are set by experience, so some work should be done to ease parameters selection problem.

Acknowledgements. The authors thank anonymous reviewers and editors for their valuable suggestions and comments on improving this paper. We also wish to thank the National Natural Science Foundation of China under Grants No. 61070090, 61003174 and 60973083), a grant from NSFC-Guangdong Joint Fund (No. U1035004), and a grant from Natural Science Foundation of Guangdong Province, China(No. 10451064101004233) for supporting this work.

References

- [1] Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification, pp. 566–581. John Wiley & Sons, New York (2001)
- [2] Brown, M., Grundyand, W., Lin, D., et al.: Knowledge-based analysis of microarray gene expression data by using support vector machines. Proceedings of the National Academy of Sciences of the United States of America 97(1), 262–267 (2000)
- [3] Isabelle, G., Jason, W., Stephen, B., Vladimir, V.: Gene selection for cancer classification using support vector machines. Machine Learning 46(1), 389–422 (2002)
- [4] Ah Hwee, T., Hong, P.: Predictive neural networks for gene expression data analysis. Neural Networks 18(3), 297–306 (2005)
- [5] Zhou, X., Liu, K.Y., Wong, T.C.: Cancer classification and prediction using logistic regression with bayesian gene selection. Journal of Biomedical Informatics 37(4), 249–259 (2004)

- [6] Li, L., Weinberg, C.R., Darden, T.A., Pedersen, L.G.: Gene selection for sample classification based on gene expression data: study of sensitivity to choice of parameters of the GA/KNN method. *Bioinformatics* 17(12), 1131–1142 (2001)
- [7] Nguyen, D.V., Rocke, D.M.: Tumor classification by partial least squares using microarray gene expression data. *Bioinformatics* 18(1), 39–50 (2002)
- [8] Huang, D.S., Zheng, C.H.: Independent component analysis based penalized discriminant method for tumor classification using gene expression data. *Bioinformatics* 22(15), 1855–1862 (2006)
- [9] Pochet, N., De Smet, F., Suykens, J.A.K., De Moor, B.L.R.: Systematic benchmarking of microarray data classification: assessing the role of non-linearity and dimensionality reduction. *Bioinformatics* 20, 3185–3195 (2004)
- [10] Schölkopf, B., Smola, A.J., Müller, K.: Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation* 10(5), 1299–1399 (1998)
- [11] Saul, L.K., Roweis, S.T.: Think globally, fit locally: unsupervised learning of low dimensional manifolds. *Journal of Machine Learning Research* 4, 119–155 (2003)
- [12] Li, B., Zheng, C.H., Huang, D.S., et al.: Gene Expression Data Classification Using Locally Linear Discriminant Embedding. *Computers in Biology and Medicine* 40, 802–810 (2010)
- [13] Klein, D., Kamvar, S.D., Manning, C.: From instance-level constraints to space-level constraints: Making the most of prior knowledge in data clustering. In: Sammut, C., Hoffmann, A.G. (eds.) *Proc. of the 19th Int'l Conf. on Machine Learning*, pp. 307–314. Morgan Kaufmann Publishers, San Francisco (2002)
- [14] Zhu, X.: *Semi-Supervised Learning Literature Survey*(EB/OL) (July 19, 2008), http://www.cs.wisc.edu/~jerryzhu/pub/ssl_survey.pdf
- [15] Zhang, D.Q., Zhou, Z.H., Chen, S.C.: Semi-Supervised Dimensionality Reduction. In: *Proc. of the 7th SIAM International Conference on Data Mining*, pp. 629–634 (2007)
- [16] Cevikalp, H., Verbeek, J., Jurie, F., Klaser, A.: Semi-supervised dimensionality reduction using pairwise equivalence constraints. In: *Proceedings of the 3rd International Conference on Computer Vision Theory and Applications* (2008)
- [17] Wei, J., Peng, H.: Neighbourhood Preserving based Semi-Supervised Dimensionality Reduction. *Electronics Letters* 44(20), 1190–1192 (2008)
- [18] Roweis, S.T., Saul, L.K.: Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science* 290(22), 2323–2327 (2000)
- [19] Chung, F.R.K.: Spectral graph theory. In: *Regional Conference Series in Mathematics* (1997)
- [20] Martinez, A.M., Kak, A.C.: PCA Versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23(2), 228–233 (2001)

A Layered Model of Artificial Emotion Merging with Attitude

Qijun Luo, Ang Zhao, and Hongxiang Zhang

Robotics Institute, Civil Aviation University of China, 300300, Tianjin, China
qjluo@cauc.edu.cn, zha_stef@sina.com, hgd_hxiang@163.com

Abstract. Service robot should have human emotions, and farthest resemble the real-life interaction when interact with humans. A layered model of artificial emotion called AME (Attitude Mood Emotion) model, merging with attitude is proposed, which is more human, and an emotion-stimulation generation mechanism based on the OCC cognitive appraisal model and the PAD emotion space is built. The concept of "Mood Baseline" is introduced to describe the interaction among emotions, mood and attitude. The validity and practicability of this model can be verified based on "FuNiu" robot.

Keywords: human-robot interaction, affective modeling, attitude.

1 Introduction

As robots become more prevalent in social spaces, and remarkable progress have been made in such as the elderly healthcare, assistant therapy of childhood autism etc [1]. People need to be able to interact with these robots in a more smooth and human way.

Early in 1995, professor R.Picard from the MIT first proposed the concept of "Affective Computing"[2]. Affective modeling is an approach of producing artificial affection, and is the basis of harmonious human-robot interaction. There are some typical models. OCC affection model [3] by Ortony and others, sorted out 22 emotion types based on cognitive mechanism, and make affect as the result of situation; in the study of affective modeling, Moshkina proposed TAME [4] model according to actions of personality, attitude, mood and emotion on affect, although TAME takes attitude into consideration, the reciprocity between different affects is not mentioned, not how humans influence robot's affect either; an affect model based on emotion-mood-attitude [5] is come up by Kirby and Simmons, which take attitude into account, but the evaluation of emotion stimulation is artificial given, with absence of cognitive appraisal mechanism. Robots need to have integrative and human like emotion model, and in human interaction, attitude takes important part. In human interaction, people take different attitude to communicate with the opposing part according to their intimate and fond extent, and this will influence the emotion expression and behavior decision. Meanwhile, according to the communication effect, people will constantly look back on their conversation attitude, so, the attitude to conversation object is changeable.

Therefore, taking attitude’s influence on emotion into account, a layered affect model called AME model is proposed, which divides affect into attitude, mood and emotion three layers. The emotion layer adopts OCC model to cognize and appraisal outside information, then map it to PAD(Pleasure Arousal Dominance) space, and the emotion stimulation vector is constructed; the mood layer which is based on “mood Baseline” is fluctuating along with the emotion and attitude; the attitude layer is determined by the intimate and fond extent between the robot and its communication partner.

2 Construction of AME Model

The overall framework of AME is shown in Fig 1. Vision, voice, touch screen and so on give the robot outside emotion stimulation as inputs, and they will be appraisal through the OCC cognitive evaluation model, after mapped to PAD emotion space to decide their type and intensity, they will act on the emotion layer. Based on the designed mood baseline, they also act on the mood layer. In the interaction with different objects, according to the past interaction time and effect after identity recognition, familiarity and fondness extent will be computed, which will act on the attitude layer.

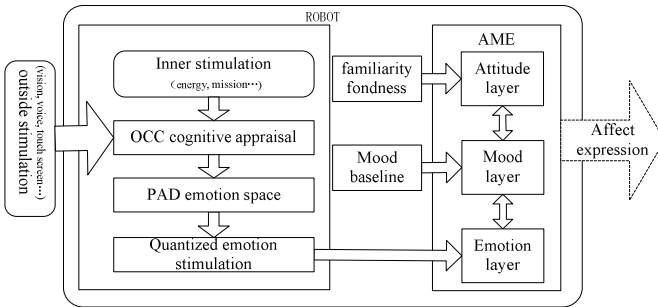


Fig. 1. Framework of AME model

2.1 Construction of Attitude Layer

Attitude is the core of AME model, which shows people’s fondness extent to different people(things). Let A represent attitude.

The attitude layer is composed of two parts: the familiarity between robot and interaction object A_f ; the fondness to interaction object A_m . familiarity A_f is computed by the interaction record, shown in formula (1).

$$A_f = \frac{1}{2} \times \left(\frac{1}{T} \times \min(\text{hours}, T) + \frac{1}{C} \times \min(\text{count}, C) \right) \tag{1}$$

Hours is the interaction time, *count* is the interaction times. T, C are constants.

Fondness A_m is decided by A_f and the past interaction effect, shown in formula (2). ΔM is the mood variation through the interaction.

$$A_m' = A_m + \Delta M \times (1 - A_f) \quad (2)$$

2.2 Construction of Mood Layer

Mood is an affect state that can influence people's whole psychology, and it has the characteristic of diffuse. Let M show the mood state.

Divide mood into negativity, positivity and calmness, and the concept of "Baseline" is introduced, which reflect people's basic mood intensity. "Baseline" is decided by the past and current incidents. If an incident happens at the d_0 day, and the mood intensity is m , the relative mood intensity changes following formula (3).

$$M_b = m - \frac{m}{1 + e^{-|d-d_0|+|10m|}} \quad (3)$$

Where, d is the number of days.

In order to limit mood intensity between -1 and +1, introduce formula (4).

$$M_d = \begin{cases} \frac{1+M_b}{1+e^{-B(M_o-C(1+M_b))}} - 1 & \text{if } M_o < 0 \\ M_b & \text{if } M_o = 0 \\ \frac{1-M_b}{1+e^{-B(M_o-C(1-M_b))}} + M_b & \text{if } M_o > 0 \end{cases} \quad (4)$$

Where M_o is the deviation intensity from baseline, M_d is the demonstrated intensity. B and C represent the change speed, decided by robot's personality.

2.3 Construction of Emotion Layer

Emotion is people's immediate emotional response to development of events or other external stimuli. Let E say with emotion.

Reference to Ekman's basic emotional classification [6], four kinds of emotions: happiness, sadness, calm and anger are built, whose strength is represented by real numbers between 0 and 1. External stimulation is cognitively appraised through the OCC cognitive model[3]. OCC model assumes that the evaluation depends on three components: events, behaviors and objects.

The appraised results from OCC model is mapped to the PAD space to quantify. The PAD affect model [7] is proposed by Mehrabian and Russell. Meanwhile, He presented the measurement tool for affect-PAD affect quantitative table, which summarizes the relationship of PAD space and emotion, shown in table 1.

Table 1. The relationship of PAD and emotion

PAD space	Emotion	PAD space	Emotion
+P+A+D	happiness	-P-A-D	sadness
+P+A-D	dependence	-P-A+D	arrogance
+P-A+D	calm	-P+A-D	anxiety
+P-A-D	mildness	-P+A+D	rage

Assume the current emotion stimulation is:

$$E = [E_p \ E_A \ E_D], \text{ Where, } E_p, E_A, E_D \in [-1, 1]$$

$$E_{intensity} = \sqrt{E_p^2 + E_A^2 + E_D^2}$$

According to reference [8], set the OCC parameters shown in table 2. S,M,L represent the influence extent of small, median and large respectively.

Table 2. The parameter design of OCC evaluation model

characters		P(Pleasure)			A(Activeness)			D(Dominance)		
		S	M	L	S	M	L	S	M	L
events	expectance			√			√			√
	not expectance	√				√		√		
	anticipated		√			√			√	
	not anticipated	√					√	√		
behaviors	self	praised			√		√			√
		criticized	√				√			√
	someone else	beneficial			√		√			√
		harmful	√				√			√
objects	familiar			√		√				√
	unfamiliar			√		√			√	
	like				√		√		√	
	dislike		√			√			√	

2.4 Reciprocity of Layers

Reciprocity between Attitude and Mood. The robot’s attitude to object influences the mood directly. When the interaction is started, the robot’s attitude to object will be reflected by M_o -the deviation intensity from baseline, shown in formula (5).

$$M'_o = (M_o + A_m) / 2 \tag{5}$$

At the end of the interaction, $M'_o = (M_{before} + M_{current})/2$. After the interaction, the mood intensity will lessens gradually to the baseline value.

Reciprocity between Emotion and Mood. According to the assigned emotion type, assume that the emotion stimulation vector composed of four emotion stimuli is:

$$E_s = [E_h \quad E_s \quad E_c \quad E_a]$$

Mood intensity is constructed by positive mood and negative mood, the mood vector is:

$$M = [M_p \quad M_n]$$

Assume the mapping matrix from emotion to mood is C , and:

$$M = E_s C$$

For the convenience of the problem explanation, assume that the robot is influenced by only one emotion stimulation at one moment, so we can assume that:

$$C = \begin{bmatrix} \lambda_{ph} & 0 \\ 0 & \lambda_{ns} \\ \lambda_{pc} & 0 \\ 0 & \lambda_{na} \end{bmatrix} \tag{6}$$

Where, λ_{ph} , λ_{pc} , λ_{ns} , λ_{na} are weights, which represent influences of happiness and calm to positive mood, and influences of sadness and anger to negative mood respectively.

The current emotion intensity is influenced by the emotion intensity and mood intensity of last moment, shown in formula (7):

$$E' = E(1 + \alpha v M) \tag{7}$$

- E' —the current emotion stimuli intensity
- E —the emotion intensity of last moment
- M —the mood intensity of last moment

When $E \geq 0$, $v = 1$, and when $E < 0$, $v = -1$, representing enhanced or weakening effect. α is constant, deciding the enhanced or weakening intensity.

Mood acts on emotion, and emotion influences mood inversely, also, the current mood is influenced by emotion and mood of last moment. Study of Rook[9] shows: the positive outside stimulation enhance the positive mood intensity, but not so much to negative mood intensity; negative outside stimulation has great effect on both kinds of mood. Based on this, let:

$$M' = \beta E$$

β is decided by the current mood intensity and emotion stimulation intensity:

$$\beta = \begin{cases} \theta_1 & M < 0 \text{ and } E > 0 \\ \theta_2 & \text{otherwise} \end{cases} \tag{8}$$

Where, $\theta_1 < \theta_2$, and can be regulated according to the robot's personality.

When the stimulation disappeared, mood will reduce to the baseline value gradually. From Picard[2], people's emotion changing law is just like the bell effect, use formula (9) to describe mood's reduction:

$$M'_d = (M_d - M_b) \times e^{-\gamma t} + M_b \quad (9)$$

M'_d —mood intensity after reduction

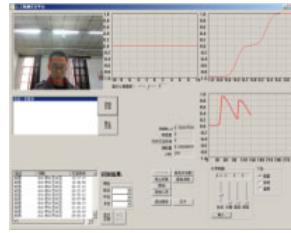
M_d —mood intensity before the reduction

γ —reduction parameter, representing the robot's personality.

3 Experiments and Analysis

3.1 Introduction of Experiment Platform

To testify the practicality and validity of AME model, "FuNiu" robot is used as experiment platform, shown in Fig.2(a), which works at airport terminal.



(a) "FuNiu" robot named LeLe (b) Interaction platform of AME model

Fig. 2. Artificial emotion interaction experiment platform

3.2 Experiment Design

The interactive platform is designed in VC++6.0, shown in Fig 2(b). Experiments are designed in the thought of scene. When interacting with passengers, the robot will first identify the passenger through face recognition based on SIFT feature, which has high reliability. Passengers can making an inquiry or give their assessment through the touch screen or voice, and the robot will serve passengers through the TTS voice output system or touch-screen prompts.

Scene 1: To testify the robot's affective change without attitude influence. Firstly, remove the attitude's influence to affect model, set the baseline value 0, let passenger A who is familiar with the robot and their former interaction effect is good interact with the robot twice. In the first interaction, passenger A gives the robot praise, and the interaction effect is relatively good; in the second interaction passenger A criticize the robot, and the interaction effect is poor.

Scene 2: Assume there are A, B two passengers, who are both familiar with the robot, and A had a good time with the robot, while B is not. Set the robot's mood baseline value -0.1. Passenger A and passenger B interact with the robot successively. After accomplishing passenger A's assignment, although having an enmity against

passenger B, the robot still accomplish B’s assignment, after that passenger B speaks highly of the robot, and the robot’s attitude to passenger B is improved.

3.3 Experiment Results and Analysis

Selecting experimental parameters. According to the experimental design and results of repeated experiments, the parameters are chosen as follows, and in specific applications, they can be adjusted as needed:

- In formula (6), $[\lambda_{ph}, \lambda_{pc}, \lambda_{ms}, \lambda_{na}] = [0.7, 0.3, 0.6, 0.4]$;
- In formula (7), $\alpha = 0.25$;
- In formula (8), $\theta_1 = 0.1, \theta_2 = 0.2$.

Analysis of experiment results. The robot’s mood changing curve of scene 1 is shown in Fig.3(a).

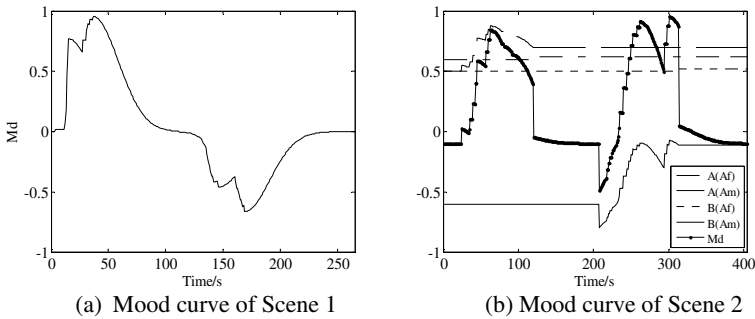


Fig. 3. Mood changing curve

$0 \leq t \leq 3$: Before passenger A’s coming, the robot is in the calm state. At the 3th second, passenger A become interact with the robot, as the attitude’s influence is eliminated, the robot is still in calm state.

$3 < t \leq 111$: At the 4th second, passenger A do the information query and other operations, as the robot accomplish all its tasks, passenger A praises the robot, and the robot’s mood intensity comes to 0.9204 at the 37th second.

$111 < t \leq 265$: At the 111th second, passenger A become do some operations and give evaluation according to the robot’s performance. The robot’s mood fluctuates, decreasing to -0.6645 at second 169. At second 200, passenger A leaves the robot, and the robot’s mood continues decrease with the original law.

Experiment results of scene 1 shows that removing the affective impact of attitude, the robot’s mood changes independent of interactive objects, and the mood changes of the robot is totally different with human - human interaction.

In scene 2, curves of the robot’s mood and attitude to interactive objects are shown in Fig.3(b).

$0 \leq t \leq 25$: Before passenger A’s arrival, the robot is in the negative state. At the 25th second, A become to interact with the robot, and their familiarity $A_f = 0.6$, fondness $A_m = 0.5$. The robot’s mood intensity increases to 0.0242.

$25 < t \leq 120$: Passenger A become do some operations, and the robot accomplishes its tasks, so A praises the robot, then the mood intensity increases gradually, to 0.8419 at second 63.

$120 < t \leq 185$: At second 120, A leaves the robot, and the mood intensity decreases to baseline value at second 185.

$185 < t \leq 208$: At second 208, passenger B becomes to interact with the robot, and their familiarity $A_f = 0.5$, fondness $A_m = -0.6$, representing that they are familiar, but the robot doesn't very like B. The mood intensity decreases to -0.4767.

$208 < t \leq 313$: The robot accomplish tasks from B, B give the robot very high marks, so the mood intensity increases gradually, to 0.9460 at second 304.

$t > 313$: At the 313th second, passenger B leaves, the mood intensity goes back to the baseline gradually.

Attitude changes to passenger A: as the interactive effect is fine, the robot's familiarity and fondness to A increases from 0.62 to 0.70.

Attitude changes to passenger B: as the robot is very happy through their interaction, at the end, $A_f = 0.52$, $A_m = -0.1$, representing there is improvement of robot's attitude to B.

The experiment results of the two scenes shows: Introduction of attitude makes the affective model more anthropomorphic, and the human-robotic interaction more in line with the human interaction law.

4 Conclusions

A layered model of artificial emotion called AME model, merging with attitude is proposed. Definition of "Baseline" is introduced to describe robot's basic mood intensity, meanwhile, OCC model and PAD space are adopted to appraisal and quantify outside stimulation. The robot's attitude to different objects is integrated to improve the anthropomorphic level of robot affect, and thus make the robot service more user-friendly. This model provides foundation for the robot affect generation and updating mechanism in human-robotic interaction.

Acknowledgments. This work was sponsored by the Fundamental Research Funds for the Central Universities (XH2009B002) and the Research Funds of Civil Aviation University of China (08CAUC-E07).

References

1. Barakova, E.I., Lourens, T.: Expressing and interpreting emotional movements in social games with robots. *Pers. Ubiquit. Comput.* 14, 457–467 (2010)
2. Picard, R.W.: *Affective computing*. MIT Press, London (1997)
3. Ortony, A., Clore, G.L., Collins, A.: *The Cognitive Structure of Emotions*, pp. 123–127. Cambridge University Press, Cambridge (1988)
4. Moshkina, L., Arkin, R.C.: On TAMEing robots. In: *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3949–3956 (2003)
5. Kirby, R., Forlizzi, J., Simmons, R.: Affective social robots. *Robotics and Autonomous Systems* 58, 322–332 (2010)

6. Ekman, P.: *Emotion in the Human Face*. University of Cambridge Press, Cambridge (1982)
7. Mehrabian, A.: Pleasure-Arousal-Dominance: a general framework for Describing and measuring individual differences in temperament. *Current Psychology: Developmental, Learning, Personality, Social* 4(14), 261–292 (1996)
8. Fang, Y., Chen, Z.-Q., Yuan, Z.-Z.: An Emotion Model Based on Artificial Intelligence. *Information and Control* 35(6), 673–678 (2006)
9. Rook, K.S.: Emotional health and positive versus negative social exchanges: A daily diary analysis. *Applied Developmental Sciences* 5(2), 86–97 (2001)

Fast Extraction Strategy of Support Vector Machines

Wei Wu, Qiang Yang, and Wenjun Yan

College of Electrical Engineering, Zhejiang University, Hangzhou, P.R. China
wj.yan@126.com

Abstract. As a universally accepted tool of machine learning, support vector machine (SVM) is efficient in most scenarios but often suffers from prohibitive complexity in dealing with large-scale classification problems in terms of computation time and storage space. To address such intractability, this paper presents a group and nearest neighbor strategy aiming to extract support vectors from training samples for obtaining the discriminant function in a fast fashion as only the support vectors contribute to the function. For non-linear cases, kernel function is investigated and adopted in this approach. The proposed strategy is described through mathematical analysis and evaluated by a set of numerical experiments. The result demonstrates that the suggested approach is effective in addressing the large-scale classification problems with acceptable complexity.

Keywords: Support Vector Machine, Group, Nearest neighbor, Pattern recognition.

1 Introduction

Support Vector Machine (SVM) is originally proposed by in [1] and now becomes one of most widely used tools in the research area of machine learning. SVM has been adopted in pattern recognition, function approximation and related areas (e.g. [2, 3]) due to its generalization capability in addressing non-linear problems.

However, it is known that though SVM provides excellent performance in many aspects, the complexity in terms of both computation time and space is still prohibitive during its application, in particular for solving large-scale problems. To significantly improve the learning speed, lots of research effort has been made and a collection of efficient methods are introduced to SVM. The majority of the existing solutions are based on various optimization algorithms and sample pretreatment, e.g. Least Squire SVM (LS-SVM) [4], Sequential Minimal Optimization (SMO) [5], Lagrangian Support Vector Machines (LSVMs) [6], Core Vector Machine (CVM) [7], and Clustering-Based SVM (CB-SVM) [8].

Sample pretreatment always works better than algorithm optimization in terms of reducing computing load, but still with long training time. Based on such recognition, this paper exploits a novel approach of fast extraction strategy of SVM which attempts to reduce the computation load of SVM so as to enhance the optimization performance. In general, the standard SVM learning problem is in fact a convex quadratic programming problem with a set of constraints [9]. The optimization of

SVM problem depends merely on the support vectors (SVs), which build up the optimal classifier. Based on such characteristic, a novel fast SVM algorithm is proposed and evaluated in this paper. The rest of this paper is organized as follows: section 2 firstly overviews the support vector machine and describes its characteristics; the selection of support vectors are discussed in section 3; followed by the numerical experiments with a set of key results and analysis; finally section 4 concludes the paper with some conclusive remarks.

2 Support Vector Machine

In literature, SVM has been proved as an efficient method for pattern recognition in a variety of applications. In this paper, we focus on the two-class pattern recognition problems. Of course, there are also many algorithms proposed to solve multi-class problems.

Given the training samples, denoted as $(x_i, y_i): x_i \in R^d, y_i \in \{-1, 1\}, i = 1, \dots, l$, where d and l are the dimension of x and the number of training samples respectively. We firstly define the samples by using $y_i = 1$ and $y_i = -1$ to indicate positive class and negative class. As a result, the two-class pattern recognition problems can be effectively transformed into a primal problem for finding the hyperplane which can be expressed as follows:

$$w^T x + b = 1 \quad (1)$$

where w is a d -dimension normal vector such that the samples can be separated with the maximum separating margin and this hyperplane is also known as the optimal separating hyperplane. Based on the risk minimization principle, the primal problem can be expressed as the following quadratic optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s.t.} \quad & y_i (w^T x_i + b) \geq 1 - \xi_i \end{aligned} \quad (2)$$

The denotations are defined as follows: C is the penalty factor reflecting the weight of classification error; ξ_i is the slack variable; w is the weighted vector determining the optimal hyperplane direction and b is the bias value.

The use of Lagrange multiplier techniques can effectively transform the problem into the dual optimization problem which can be described as follows:

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{s.t.} \quad & \sum_{i=1}^l \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C, \\ & i = 1, \dots, l \end{aligned} \quad (3)$$

The discriminant function takes the following form

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^l \alpha_i^* y_i (x_i^T x) + b^* \right\} \quad (4)$$

where (α^*, b^*) is the optimal solution of the quadratic problem.

For the non-linear problems, we can choose appropriate function $\phi(x)$ to map samples to high-dimension space where the samples could be linearly separable. Here, we define the kernel function $K(x, x') = \langle \phi(x), \phi(x') \rangle$, where $\langle \cdot, \cdot \rangle$ represents the inner product. If the kernel function could meet the Mercer Condition, $x \cdot x'$ in the dual optimization problem and discriminant function can be replaced by $K(x, x')$. In this paper, we choose the RBF kernel function which can be expressed as follows: $K(x_i, x_j) = \exp(-\|x_i - x_j\|/\sigma^2)$.

Through the adoption of this approach, the pattern recognition for non-linear cases can be greatly simplified.

Consider the KKT condition for SVM given in [10]:

$$\begin{cases} \alpha_i = 0 \Rightarrow y_i f_i \geq 1 \\ \alpha_i = C \Rightarrow y_i f_i \leq 1 \\ 0 < \alpha_i < C \Rightarrow y_i f_i = 1 \end{cases} \quad (5)$$

It's apparent that only the support vectors on and near the separating hyperplane can be corresponding to $\alpha_i \neq 0$. In other words, the other samples have no impact on the machine training process. These support vectors are often represent a small portion of the overall samples. If these support vectors can be appropriately selected in advance, the training process can be significantly simplified with the equivalent performance to the case of using all the samples.

3 The Selection of Support Vectors

To obtain support vectors, we must distinguish the boundary samples exactly. In this paper, a novel Group and Nearest Neighbor SVM (GNNSVM) is proposed which can be described as follows: (1) find cluster centers for both of the classes in this two-class pattern recognition problem; (2) divide the whole samples into several groups by computing the distance between the samples and the centers and delete the groups that include no support vectors; (3) select candidate support vectors in the remaining groups by nearest neighbor strategy. All the candidate support vectors are considered as support vectors though some of them are in fact not support vectors. They are used to compute the discriminant function. Extractions of support vectors for positive and negative classes are completed separately, but with the similar process, and hence we take the extraction for positive class for example in the discussion.

3.1 Cluster Center Identification

For positive samples, compute the mean $\bar{x}_- = \frac{1}{n_-} \sum_{i=1}^{n_-} x_{k_-}$, n_- is the number of negative samples, x_{k_-} is the negative sample. The sample near \bar{x}_- most is the center of negative class.

3.2 Divide Samples into Groups

Define the distance between samples x_i and x_j as $D_{i,j} = \|x_i - x_j\|$. In non-linear problems, define the distance as

$$D_{i,j} = \|\phi(x_i) - \phi(x_j)\| = \sqrt{K(x_i, x_i) - 2K(x_i, x_j) + K(x_j, x_j)} \quad (6)$$

Compute the distances between all the samples and \bar{x}_- , and call them negative center distance. Put the samples in order according to negative center distance and divide samples into L_- groups averagely. The former group consists of samples with shorter distances than the later group.

To reduce the computing load, select T continuous groups from obtained groups that include all the support vectors:

Mark the group including the positive sample with shortest negative center distance as the first support group. Mark the group including the negative sample with longest negative center distance as the candidate last support group. If the candidate last support group is the last group of all L_- groups, mark it as the last support group, otherwise select the next group as the last support group.

It's not difficult to see that the other samples (not included in these T groups) couldn't be the support vectors. There is an exceptional case that no group consists of both positive and negative samples. If so, the group of negative samples with longest center distance and the group of positive samples with shortest center distance constitute the support groups. The samples in these two groups constitute candidate support vectors.

3.3 Select Candidate Support Vectors by Nearest Neighbor Strategy

Consider the selected T groups and given that the positive sample x_i in these groups belongs to support group t , $t \in \{1, \dots, T\}$, if the sample meets the following criteria, we consider it as a candidate support vector:

- 1) x_i belongs to the first support group;
- 2) otherwise, select m samples in the former support group with shortest distance to x_i , and the number of negative samples in these m samples is more than M .

After above steps, all the candidate positive support vectors are selected. Similarly, all the candidate negative support vectors can be selected.

Few support vectors will escape from candidate support vectors. There are also some candidate support vectors are not support vectors. Setting appropriate parameters such as L_* , m and M can make candidate support vectors in proper scale.

Increasing L_* can reduce the computing scale while the too large L_* will make the separating performance bad. Large m will increase the computing scale but can improve precision of support vector judgment. Large M will make a good precision of support vector judgment but some support vectors may be omitted.

4 Numerical Experiment and Key Result

In this section, the standard datasets are used in the numerical experiments to evaluate the proposed GNNSVM solution by using the standard SVM as a comparative benchmark. In the experiments, we adopt three different UCI datasets: - Cylinder, Credit and Tic-Tac-Toe datasets. For each dataset, half samples are used for training and the others for algorithm performance test. Take 10 trials for each dataset. The experiments are carried out in the platform with the following configurations: CPU-Pentium 4, Memory-1.0GRAM, and OS-Windows XP. The algorithm is implemented using Matlab 7.0.

Table 1. Standard SVM and GNNSVM

Dataset	Standard SVM			GNNSVM		
	Time/s	SVs	TA/%	Time/s	SVs	TA/%
Cylinder	100.6	93	85.7	36.2	129	83.9
Credit	119.9	98	92.2	40.3	147	90.6
Tic-Tac-Toe	543.8	159	97.1	230.1	222	95.7

Table 1 shows the result obtained from the numerical experiments which indicates that the proposed GNNSVM can significantly reduce training time with little precision degradation.

5 Conclusions and Future Work

In this paper, we present a novel GNNSVM approach to significantly reduce the computation complexity in both time and space in coping with the large-scale SVM problems. In this suggested solution, group and nearest neighbor method are used to extract support vectors efficiently and fast. Experiment results have proved the proposed strategy reasonable and effective. There are some work needs further research effort based on the outcome of this paper, such as the selection of support vectors through a number of iterations and the parameter optimization in GNN strategy. The further research will be reported in the future publications.

Acknowledgements. This work is support in part by the National High Technology Research and Development Program of China (“863” Program) under the grant No.2007AA05Z232.

References

1. Vapnik, V.N.: The nature of statistical learning theory. Springer, New York (1995)
2. Begg, R.K., Palaniswami, M., Owen, B.: Support vector machines for automated gait classification. *IEEE Trans. Biomedical Engineering* 52(5), 828–838 (2005)
3. He, P., Wang, Y.L., Gui, W.H., Kong, L.S.: Chaotic time series analysis and SVM prediction of alumina silicon slag composition. In: *Chinese Control Conference, China*, pp. 1273–1277 (2010)
4. Suykens, J.A.K., Vandewalle, J.: Least squares support vector machine classifiers. *Neural Process Letter* 9(3), 293–300 (1999)
5. Platt, J.C.: Fast training of support vector machines using sequential minimal optimization. In: *Advances in Kernel Methods-Support Vector Learning*, pp. 185–208. MIT Press, Cambridge (1998)
6. Mangasarian, O.L., Musicant, D.R.: Lagrangian support vector machines. *Journal of Machine Learning Research* 1, 121–177 (2001)
7. Tsing, I.W., Kwok, J.T., Cheung, P.M.: Core vector machines: fast svm training on very large data sets. *Journal of Machine Learning Research* 6(4), 363–392 (2005)
8. Yu, H., Yang, J., Han, J.W.: Making SVMs scalable to large data sets using hierarchical cluster indexing. *Data Mining and Knowledge Discovery* 11(3), 295–321 (2005)
9. Vapnik, V.N.: An overview of statistical learning theory. *IEEE Trans. on Neural Networks* 10(5), 988–999 (1999)
10. Meng, D.Y., Xu, Z.B., Jing, W.F.: A more efficient preprocessing method for support vector classification. In: *Proc. of Int. Conf. on Neural Networks and Brain, Beijing*, pp. 1173–1177 (2005)

A Local Feature Selection Approach for Clustering

Bing Gui

Department of Computer Science, Shanghai Institute of Technology,
Shanghai 201418, China
Guib@sit.edu.cn

Abstract. A local feature saliency approach is introduced which locally selects relevant features for particular clusters. In addition, the original feature saliency algorithm is improved to deal with redundant features by using local correlations. Experiments using 10 benchmark data sets are conducted for estimating the performances of proposed approach. For comparison, results obtained from the traditional algorithms are also provided. The experimental results show that the proposed local feature selection approach outperforms the traditional algorithms.

Keywords: Clustering, local correlation, local feature saliency, local feature selection.

1 Motivation and Preliminary

The fundamental motivation for feature selection is the curse of dimensionality [1]. Data points which are close to each other in a low dimensional feature space are likely to be distant in a high dimensional feature space. This makes it difficult to partition data or a feature space into a set of classes and to assign a data point to one of the classes in both supervised and unsupervised learning. Moreover, noisy features which do not contribute to the learning process often degrade the learning performance.

Generally, the selection of features can be divided into two categories: filter and wrapper. Law [2] proposed a wrapper approach for unsupervised feature selection in which a real-valued quantity called feature saliency is estimated for the relevance of each feature to the unsupervised learning process. Based on the idea of feature saliency, this paper proposes a new approach using local feature selection in unsupervised learning to find embedded clusters with local feature saliency. The proposed approach provides each cluster with a different subset of features during the learning process. Law [2] assumes that features are independently modeled with feature saliency. Under this assumption, a redundant feature cannot be removed because its distribution is independent of the component label given by another feature. Therefore, a local feature redundancy removal approach is also proposed.

This paper is organized as follows. Section 2 gives overviews of clustering and feature selection approaches. Section 3 introduces the proposed feature selection approach with local feature saliency and redundancy removal methods. Section 4 empirically evaluates the approach using ten benchmark data sets. Finally, section 5 summarizes the paper.

2 Survey of Clustering and Feature Selection Approaches

As mentioned above, the existence of irrelevant and redundant features can degrade unsupervised learning performance. Therefore, unsupervised feature selection approaches are needed.

2.1 Clustering Algorithms for Unsupervised Learning

Generally, clustering algorithms are divided into two categories: partitional and hierarchical. Many clustering algorithms have been proposed within these two categories [3], [4]. This paper mainly focuses on partitional clustering using a finite mixture model (FMM).

In a FMM, data points are assumed to be generated from a mixture of K density functions, in which K is generally regarded as the cluster number. The probability density function with K clusters is given as follows

$$p(\mathcal{Y}) = \sum_{k=1}^K \pi_k p(y_i | x_i = k, \theta_k), \quad (1)$$

where π_k is the mixing probability of each mixture component. The log-likelihood of a mixture model is described as

$$L = \sum_{i=1}^N \log \sum_{k=1}^K \pi_k p(y_i | x_i = k, \theta_k). \quad (2)$$

Direct maximization of likelihood in equation (2) is a difficult task. Therefore, the expectation maximization (EM) algorithm is utilized to find a local maximum. The E and M steps are iterated until a convergence criterion is met. In clustering, the E and M steps can be considered as the estimation of data membership (posterior) and mixture model parameters respectively.

2.2 Wrapper Approaches for Feature Selection

Wrapper approaches incorporate the learning process into feature selection. Generally, wrapper approaches consist of three components. Firstly, a search strategy needs to be determined. An exhaustive search would definitely find the optimum feature subset which, however, is computationally impractical. Sequential forward and backward search are developed in a greedy manner which only finds local optimality [5]. In order to escape from local optimum, random search methods such as genetic algorithms [6] and random mutation hill climbing [7] have been proposed. After determining the search strategy, one of the clustering algorithms is applied. Finally, to select relevant features during learning process, a feature evaluation criterion must be selected. Maximum likelihood is a prevalent evaluation criterion. Maximum likelihood selects features which can be best modeled as a FMM. It was discovered that maximum likelihood criterion do not monotonically change along with cluster assignments in different feature subspaces. Based on this property, Gennari [8] proposed a feature attention approach which looked for features that mostly influence the clustering decision. The idea of feature attention was incorporated into the FMM approach as a latent variable (feature saliency) that measures the feature relevance.

3 Proposed Local Feature Selection in Unsupervised Learning

In this section, a simultaneous local feature selection and clustering approach is proposed, based on local feature saliency. Maximizing the likelihood of a FMM is performed for clustering by means of an EM algorithm. The widely used Gaussian distribution is considered as the probability density functions (pdf) in the FMM. Apart from the E and M steps in the EM algorithm, a new step (the T step) is incorporated to remove redundancies based on the considerations of local and global correlation between features.

3.1 FMM with Local Feature Saliency

This section introduces the concept of local feature saliency and derives an EM algorithm for clustering. To incorporate local feature saliency into a FMM, features are assumed to be independent given the hidden label. Equation (1) therefore becomes

$$p(\mathcal{Y}) = \sum_{k=1}^K \pi_k \prod_{l=1}^D p(y_{il}|x_i = k, \theta_{kl}), \quad (3)$$

where $p(y_{il}|x_i = k, \theta_{kl})$ is the pdf of the l th feature in the k th cluster. Although the assumption of feature independency enables the derivation of the EM algorithm of local feature saliency, it also brings the problem of removing redundant features which will be discussed in section 3.3.

Feature saliency focuses on removing irrelevant features. The definition of irrelevant features introduced by [9] has been adopted in traditional global feature saliency: *the l th feature is irrelevant if its distribution is independent of the class labels*. The pdf of an irrelevant feature should follow a common density function for all clusters. In terms of local feature saliency, the l th feature is irrelevant for the k th cluster if its distribution is common with other clusters in the l th feature space.

The pdf of a locally irrelevant feature is denoted by $q(y_{il}|x_i = k, \lambda_i^{common})$, which is determined by a common pdf in a feature. Define a set of binary parameters $\Phi = (\phi_{k1}, \dots, \phi_{kD})$ such that $\phi_{kl} = 1$ if feature l is relevant to cluster k and $\phi_{kl} = 0$, otherwise, equation (3) is written as

$$p(\mathcal{Y} | \Phi) = \sum_{k=1}^K \pi_k \prod_{l=1}^D p(y_{il}|x_i = k, \theta_{kl})^{\phi_{kl}} q(y_{il}|x_i = k, \lambda_i^{common})^{1-\phi_{kl}}. \quad (4)$$

where ϕ_{kl} is a hidden parameter indicating the relevance of l th feature to k th cluster. Local feature saliency is therefore defined as $\rho_{kl} = p(\phi_{kl} = 1)$, which is the probability that l th feature is relevant to k th cluster. Equation (4) with local feature saliency therefore takes the following form

$$p(\mathcal{Y}) = \sum_{k=1}^K \pi_k \prod_{l=1}^D \rho_{kl} p(y_{il}|x_i = k, \theta_{kl}) + (1 - \rho_{kl}) q(y_{il}|x_i = k, \lambda_i^{common}), \quad (5)$$

where $\{\pi, \rho, \theta, \lambda\}$ are parameters that need to be maximized in the M step. The transformation from equation (4) to (5) is because parameter ϕ_{kl} is binary.

The FMM with local feature saliency can be considered as a competitive learning algorithm. Salient features contributing to discriminate clusters should be modeled by

a probability $p(y_{il}|x_i = k, \theta_{kl})$, providing higher likelihood. On the other hand, non-salient features are commonly distributed in $q(y_{il}|x_i = k, \lambda_l^{common})$, which is more competitive in maximizing likelihood.

3.2 Local Feature Saliency Estimation by EM Algorithm

Taking parameter ϕ_{kl} and x_i as hidden variables and $\{\pi, \rho, \theta, \lambda\}$ as parameters to be maximized, the EM algorithm for local feature saliency is derived as follows.

E-step:

In the E-step, the expectation of complete log-likelihood is computed. Based on equation (5), $E(\log p(\mathcal{Y}, \Phi))$ is derived as follow

$$\begin{aligned} E(\log p(\mathcal{Y}, \Phi)) &= \sum_{i,k} p(x_i = k | y_i) \log \pi_k \\ &+ \sum_{k,l} \sum_i p(x_i = k, \phi_{kl} = 1 | y_i) (\log p(y_{il} | \theta_{kl}) + \log \rho_{kl}) \\ &+ \sum_{k,l} \sum_i p(x_i = k, \phi_{kl} = 0 | y_i) (\log q(y_{il} | \lambda_l^{common}) + \log(1 - \rho_{kl})), \end{aligned} \quad (6)$$

Using Bayes' rule, three posterior probability functions in equation (6) are induced as follows

$$p(x_i = k | y_i) = \frac{\pi_k \prod_l (\rho_{kl} p(y_{il} | \theta_{kl}) + (1 - \rho_{kl}) q(y_{il} | \lambda_l^{common}))}{\sum_k \pi_k \prod_l (\rho_{kl} p(y_{il} | \theta_{kl}) + (1 - \rho_{kl}) q(y_{il} | \lambda_l^{common}))}, \quad (7)$$

$$p(x_i = k, \phi_{kl} = 1 | y_i) = \frac{\rho_{kl} p(y_{il} | \theta_{kl})}{\rho_{kl} p(y_{il} | \theta_{kl}) + (1 - \rho_{kl}) q(y_{il} | \lambda_l^{common})} p(x_i = k | y_i), \quad (8)$$

$$p(x_i = k, \phi_{kl} = 0 | y_i) = \frac{(1 - \rho_{kl}) q(y_{il} | \lambda_l^{common})}{\rho_{kl} p(y_{il} | \theta_{kl}) + (1 - \rho_{kl}) q(y_{il} | \lambda_l^{common})} p(x_i = k | y_i), \quad (9)$$

where equation (7) indicates the membership of data point y_i to cluster k and equation (8) and (9) indicate the membership of data point y_i to cluster k when feature l is either relevant or irrelevant to cluster k . Substituting the posterior value in equations (7), (8) and (9) in equation (6), the expected complete data log-likelihood are maximized by the M-step.

M-step:

In the M-step, parameters $\{\pi, \rho, \theta, \lambda\}$ are updated by maximizing equation (6). The updated parameters are derived as follows

$$\pi_k = \frac{\sum_i p(x_i = k | y_i)}{\sum_i \sum_k p(x_i = k | y_i)}, \quad (10)$$

$$u_{\theta_{kl}} = \frac{\sum_i p(x_i = k, \phi_{kl} = 1 | y_i) y_{kl}}{\sum_i p(x_i = k, \phi_{kl} = 1 | y_i)}, \quad (11)$$

$$\sigma_{\theta_{kl}}^2 = \frac{\sum_i p(x_i=k, \phi_{kl}=1|y_i) (y_{kl} - u_{\theta_{kl}})^2}{\sum_i p(x_i=k, \phi_{kl}=1|y_i)}, \quad (12)$$

where $u_{\theta_{kl}}$ and $\sigma_{\theta_{kl}}^2$ are mean and variance in pdf $p(y_{il}|\theta_{kl})$. On the other hand, mean and variance in pdf $q(y_{il}|\lambda_l^{common})$, are derived with the *a posteriori* value in equation (9) as follows

$$u_\lambda = \frac{\sum_i (\sum_k p(x_i=k, \phi_{kl}=0|y_i)) y_{kl}}{\sum_i p(x_i=k, \phi_{kl}=0|y_i)}, \quad (13)$$

$$\sigma_\lambda^2 = \frac{\sum_i (\sum_k p(x_i=k, \phi_{kl}=0|y_i)) (y_{kl} - u_\lambda)^2}{\sum_i p(x_i=k, \phi_{kl}=0|y_i)}, \quad (14)$$

It can be observed that the *a posteriori* $p(x_i = k, \phi_{kl} = 1|y_i)$ measures the importance of the i th data point to the k th cluster when the l th feature is used. Therefore, it is natural to obtain the local feature saliency ρ_{kl} from $p(x_i = k, \phi_{kl} = 1|y_i)$ as follows

$$\rho_{kl} = \frac{\sum_i p(x_i=k, \phi_{kl}=1|y_i)}{\sum_i p(x_i=k, \phi_{kl}=1|y_i) + \sum_i p(x_i=k, \phi_{kl}=0|y_i)}, \quad (15)$$

With equation (7) to (15), clustering and local feature selection are achieved simultaneously. The two steps iterate until the algorithm converges to a FMM with local feature saliency associated with each cluster.

3.3 Proposed Approach for Redundant Feature Removal

It has been mentioned in section 3.1 above that the proposed approach models all features independently. This leads to the problem that the proposed approach cannot remove redundant features. To cope with the problem, a new step is incorporated to remove redundant features based on feature correlation.

Feature correlation is the best known measure of redundancy between two features. Generally, feature correlation is measured by a variable called the correlation coefficient as follows

$$\delta(x, y) = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}}, \quad (16)$$

where var is the variance of a feature and cov is the covariance between two features. The value of $\delta(x, y)$ equals one or minus one if x and y are completely correlated (either x or y is redundant).

In terms of local feature selection, the correlation coefficient measure needs to be considered locally. In each E and M steps, feature saliency ρ_{kl} is computed and used to define relevant features with $\rho_{kl} \geq \tau$, where τ is a threshold value. Because redundant features are categorized as salient features relevant for clustering, only correlations of salient features are investigated for redundancy removal. The local salient feature correlation coefficient for k th cluster is denoted as follows

$$\delta(l_1, l_2)^{\text{local}} = \frac{\text{cov}(l_1, l_2)^{\text{local}}}{\sqrt{\text{var}(l_1)^{\text{local}} \text{var}(l_2)^{\text{local}}}}, \quad (17)$$

where l_1 and l_2 are any of the two salient features and $\text{cov}(l_1, l_2)^{\text{local}}$ and $\text{var}(l)^{\text{local}}$ are given as follows

$$\text{cov}(l_1, l_2)^{\text{local}} = \frac{\sum_i p(x_i=k|y_i)(y_{i1}l_2 - u_{j1}l_2)(y_{i1}l_2 - u_{j1}l_2)^T}{\sum_i p(x_i=k|y_i)}, \quad (18)$$

$$\text{var}(l)^{\text{local}} = \frac{\sum_i p(x_i=k|y_i)(y_{i1} - u_{j1})(y_{i1} - u_{j1})^T}{\sum_i p(x_i=k|y_i)}, \quad (19)$$

where $u_{j1}l_2 = \frac{\sum_i p(x_i=k|y_i)y_{i1}l_2}{\sum_i p(x_i=k|y_i)}$ and $u_{j1} = \frac{\sum_i p(x_i=k|y_i)y_{i1}}{\sum_i p(x_i=k|y_i)}$.

Because $0 \leq |\delta(l_1, l_2)|^{\text{local}} \leq 1$, the degree of correlation between two features can be measured by a threshold value v , where $v \leq |\delta(l_1, l_2)|^{\text{local}} \leq 1$, denoting the fact that two features are highly correlated. Moreover, a further validation method for identifying redundant features is also proposed. As there is no label information provided in clustering learning, one of the methods for estimating clustering quality is the separability between different clusters. Denote V as the sum of the posterior probability of data as follows

$$V = \sum_{i=1}^N p(x_i = k|y_i). \quad (20)$$

Because $p(x_i = k|y_i)$ equals one if a data point indeed belongs to a cluster, a larger V value indicates better cluster separability. Keeping parameters $\{\pi, \theta, \lambda\}$ fixed, variable V can be maximized using a tuning saliency parameter ρ . By decreasing either the ρ value of two locally correlated features to a small value, the redundancy of the feature can be indicated by the value of V . If the value of V decreases, the corresponding feature with decreased ρ affects the quality of clustering, which therefore cannot be removed. By contrast, features with a decreased ρ but an increased or equal V value are regarded as redundant. The redundancy removal approach is denoted as the T step in the algorithm.

4 Experimental Results

In this section, the performances of the proposed approach were empirically evaluated using 10 benchmark data sets selected from University of California, Irvine (UCI) machine learning repository [11]. For comparison, results from the conventional EM algorithm and global feature saliency algorithm [2] were provided.

In the experiments, the EM, global feature saliency (GFS) and local feature saliency with (LFS-T) and without (LFS) the T step approaches were tested for 20 times. The results were presented as accuracies in percentages with means and standard deviations in table 1. It is clear from the results that feature saliency approaches generally

outperform the conventional EM approach. Moreover, the accuracies obtained from the local feature saliency approaches are higher than or similar to the ones obtained from the global feature saliency approach. This is because local feature saliency reveals cluster-wise feature relevance which provides more accurate information about the structure of the mixture model generating the data. In terms of the proposed approach, the T step is incorporated into the local feature selection approach and the saliency value of local redundant features are decreased. Compared to the approach without the T step, clustering accuracies are further improved in iris, mammography, texture, wine and zernike data sets and are preserved in the other ones.

Table 1. Experimental results using the benchmark data sets

Data set name	Accuracy (%)			
	EM	GFS	LFS	LFS-T
Breast-Cancer	94.28	95.02	96.05	96.05
Image segmentation	61.30	64.37	64.48±1.55	65.15
Ionosphere	71.23	65.24	67.85±0.30	67.52
Iris	90.67	90.6667	91.00±0.82	94.67
Mammography	80.82	80.12	81.81	81.93
Parkinson	66.30	64.43	64.43	65.46
Texture	57.17	63.46±0.42	64.31±1.24	65.24±0.51
Wdbc	91.20	93.50	93.67	93.67
Wine	97.19	97.19	96.43±0.56	99.25±0.57
Zernike	45.27	45.45	48.65±1.01	51.65±1.27

Table 1 also shows two exceptions in the ionosphere and parkinson data sets that the feature saliency approaches degrade the clustering performance. By using the principle component analysis (PCA), it was discovered that data samples in the two data sets were highly overlapped. The obtained clusters only show natural patterns in the data sets, which therefore are highly distorted from the actual clusters. As mentioned above in section 2, the feature saliency approach selected features that would best discriminate clusters. For the parkinson and ionosphere data sets, the selected features can only encourage to find the distorted clusters, therefore leading to worse results. Therefore, it can be concluded that the proposed approach is more appropriate for well-separated data sets.

5 Summary

This paper has presented a new approach for local feature selection. The proposed approach employs local feature saliency to measure the relevance of features for each cluster. The deficiency of being unable to detect local feature relevance in the global approach is solved. Moreover, a local correlation approach is also proposed to deal with the deficiency that redundant features cannot be removed in the global approach.

The proposed approach has been empirically tested and compared with the EM, global and local feature saliency without the T step approaches using ten benchmark data sets from the UCI repository. The results have shown that the proposed approach

are superior to the global approach, and generally outperforms the EM algorithm using the data sets without severely overlapped clusters. Further research is required to remove the preference of using well-separated data sets during feature selection. Moreover, the time complexity of the redundancy removal step can be further reduced.

Acknowledgments. This work is supported by the project (1010K116122-YJ2011-71) of Shanghai institute of technology.

References

1. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning*. Springer, New York (2001)
2. Law, M.H., Figueiredo, M.A.T., Jain, A.K.: Simultaneous feature selection and clustering using mixture models. *PAMI* 26(9), 1154–1166 (2004)
3. Berkhin, P.: Survey of clustering data mining techniques. *Grouping Mult. Dime. Data*, 25–71 (2006)
4. Pham, D.T., Afify, A.A.: Clustering techniques and their applications in engineering. *Proc. IMechE C, J. Mech. Eng. Sci.* 221, 1445–1459 (2007)
5. Kittler, J.: Feature set search algorithms. In: *Patt. Reco. and Sign. Proc.*, pp. 41–60 (1978)
6. Goldberg, D.E.: *Generic Algorithm in Search, Optimization and Machine Learning*. Kluwer Academic Publishers, Boston (1989)
7. Farmer, M.E., Bapna, S., Jain, A.K.: Large scale feature selection using modified random mutation hill climbing. In: *Proc. of the 17th Int. Conf. on Patt. Reco.*, vol. 2, pp. 287–290 (2004)
8. Gennari, J.H.: Concept formation and attention. In: *Proc. of the 13th Annu. Conf. of the Cogn. Scie. Soci.*, pp. 724–728 (1991)
9. Pudil, P., Novovicova, J., Kittler, J.: Feature selection based on the approximation of class densities by finite mixtures of the special type. *Patt. Reco.* 28(9), 1389–1398 (1995)
10. Blake, C.L., Merz, C.J.: *UCI Repository of Machine Learning Databases*. University of California, department of information and computer science, Irvine, CA (1998)
11. Ian, H.W., Eibe, F.: *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edn. Morgan Kaufmann, San Francisco (2005)

Semantics-Preserving Fusion of Structures of Probabilistic Graphical Models

Kun Yue¹, Yunlei Zhu¹, Kailin Tian², and Weiyl Liu¹

¹ Department of Computer Science and Technology, School of Information Science and Engineering, Yunnan University, Kunming, P.R. China

² Library of Southwest Forestry University, Kunming, P.R. China
kyue@ynu.edu.cn

Abstract. This paper gives an approach for fusing the directed acyclic graphs (DAGs) of BNs, an important and popular probabilistic graphical model (PGM). Considering conditional independence as the semantics implied in a BN, we focus on the DAG fusion while preserving the semantics in all participating BNs. Based on the concept and properties of Markov equivalence, we respectively give the algorithms for fusing equivalent and inequivalent common subgraphs of all participating BNs.

Keywords: Probabilistic graphical model (PGM), Bayesian network (BN), Fusion, Conditional independence, Markov equivalence.

1 Introduction

As an important and popular probabilistic graphical model (PGM), Bayesian network (BN) is the well accepted framework for representing and reasoning uncertain knowledge. BN is a directed acyclic graph (DAG) of random variables and each node has a conditional probability table (CPT) to represent the dependencies between variables [1]. It is desirable to achieve the global knowledge of multiple BNs in a certain domain and is necessary to establish an approach for fusing BNs. We know that the DAG construction is the critical task for BN construction [1], which does make sense in BN fusion [2, 3]. Therefore, we focus on DAG fusion in this paper, as the basis for BN fusion and then the ultimate uncertain knowledge fusion.

Matzkovich [4] gave the method for topological fusion of BNs by first obtaining a consensus structure and then estimating the model's parameters. Sagrado et al. [5] proposed the method for combining independence graphs based on the union and intersection of independencies. Richardson et al. [6] constructed a prior distribution over all DAGs from input BNs, disregarding the quantitative parts. Nielsen et al. [7] gave the method for fusing BNs based on formal argumentation in multi-BN based systems. Almost all these methods for fusing BN's DAGs have not considered the preserving of equivalent independencies commonly implied in all participating BNs. Looking upon the conditional independence as the semantics implied in a BN, we focus on the DAG fusion while preserving the semantics in all participating BNs.

Fortunately, Markov equivalence [8] provides a formal method to decide whether two given DAGs are equivalent w.r.t. the semantics. A partially acyclic directed graph

(PDAG) is used to represent equivalence classes of DAGs [8]. The completed PDAG (CPDAG) corresponding to an equivalence class is the PDAG consisting of a directed edge for every compelled edge and an undirected edge for every reversible edge [9]. Andersson et al. pointed out that the equivalence class of a certain DAG can be uniquely represented by the corresponding CPDAG [10]. For DAG fusion, Markov equivalence establishes the basis for determining whether the subgraphs of the nodes that are commonly included in the participating DAGs are equivalent semantically. CPDAG gives a unique representation of the equivalence class of a DAG.

In this paper, we first extract the subgraphs including common nodes in all participating DAGs. We then derive the CPDAGs respectively followed by obtaining the intersection of all these CPDAGs. Accordingly, we recover the DAG based on the PDAG's consistent extension [11]. Centered on the DAG commonly implied in the participating DAGs, we can easily add all the other edges.

2 Preliminary

The skeleton of any DAG is the undirected graph resulting from ignoring the directionality of every edge. The v -structure in a DAG G is an ordered triple of nodes (x, y, z) such that G contains the directed edges $x \rightarrow y$ and $z \rightarrow y$, and x and z are not adjacent in G [11]. The equivalence between two DAGs can be determined by Theorem 1.

Theorem 1 [11]. Two DAGs are equivalent if and only if they have the same skeletons and the same v -structures.

The completed PDAG, or CPDAG for short, corresponding to an equivalence class is the PDAG consisting of a directed edge for every compelled edge in the equivalence class, and an undirected edge for every reversible edge. A CPDAG for a given equivalence class of BN structures is unique [9]. We use G^* to denote the CPDAG of DAG G . The concept of consistent extension of a PDAG is used to obtain some certain members of the equivalence class. For a CPDAG G^* , every DAG contained in the equivalence class of G^* is a consistent extension of G^* [12].

Meek [12] gave the rule-based algorithm, DAG-To-CPDAG (say Algorithm 1), to transform a PDAG into a CPDAG, which we adopt to obtain the unique representation of independencies implied in a DAG. Chickering [11] gave the algorithm, PDAG-to-DAG (say Algorithm 2), to obtain the consistent extension of PDAGs, which we adopt to obtain the ultimate result of DAG fusion.

3 Fusing DAGs of BNs

3.1 Fusing Equivalent Common Subgraphs

Let G_1 and G_2 be two DAGs of two BNs respectively. If both G_1 and G_2 are Markov equivalent to some certain subgraphs of G_3 , then G_3 is expected to be the fusion result of G_1 and G_2 . Meanwhile, the Markov-equivalence-based representation of inclusion relationship between DAGs makes the equivalent independencies of each input DAG

be preserved in the result. Therefore, we establish our method based on this idea, such that the original conditional independencies can be preserved as much as possible.

Definition 1. We say DAG G_1 can be embedded into DAG G_2 , if G_1 is Markov equivalent to a subgraph of G_2 .

Given two DAGs $G_1=(V_1, E_1)$, $G_2=(V_2, E_2)$, if there exists a DAG $G_3=(V_3, E_3)$, such that $V_3=V_1\cup V_2$, $E_3=E_1\cup E_2$, and both G_1 and G_2 can be embedded into G_3 , then G_3 is expected to be the result. Whether the DAG, like G_3 , exist for any two given DAGs and how we can obtain it are critical. Actually, in view of semantics preserving and people's intuitions, the result DAG ought to satisfy the following properties: It should contain all independencies implied in the two participating DAGs and should not contain the independencies that are not implied in any of the two participating DAGs.

The first property means that all the v-structures in G_1 and G_2 should be included in G_3 . The second property means that there are no new v-structures and cycles generated in G_3 . This requires that the subgraph composed of all common nodes in G_1 and G_2 are Markov equivalent w.r.t. independencies, described by Theorem 2 as follows:

Theorem 2. Let G_1 and G_2 be two DAGs. There must exist a DAG G_3 that G_1 and G_2 can be embedded into if the subgraphs of all common nodes in G_1 and G_2 are Markov equivalent.

3.2 Fusing Inequivalent Common Subgraphs

However, the two properties of the result DAG are too strict for real-world situations, since the DAG in line with Theorem 2 frequently does not exist. Thus, it is necessary to fuse the given DAGs under more general cases than that in Theorem 2. If the subgraphs of common nodes are not Markov equivalent, we consider preserving the original semantics as much as possible. That is, we will obtain the DAG of common nodes preserving the most equivalent semantics.

Definition 2. [10] Let $G_1=(V_1, E_1)$ and $G_2=(V_2, E_2)$ be two DAGs. G_1 is larger than G_2 , denoted $G_1\subseteq G_2$ if $V_1\subseteq V_2$ and $E_1\subseteq E_2$.

Definition 3. The maximal equivalent common subgraph of G_1 and G_2 , written G^* , is the subgraph composed of common nodes in G_1 and G_2 , and (1) G^* can be embedded into G_1 and G_2 respectively, (2) there does not exist any other subgraphs that are larger than G^* satisfying (1).

Our basic ideas to obtain the maximal equivalent common subgraph are given in Algorithm 3.

Algorithm 3. Max-Subgraph

Input: $CG^*=(CV_1\cap CV_2, CE_1\cap CE_2)$

Output: Maximal extensible subgraph G_M^*

Steps:

IF CG^* is a DAG THEN RETURN CG^*

Initialization:

$V\leftarrow CV_1\cap CV_2$, $E\leftarrow CE_1\cap CE_2$, $G=(V, E)$

Initialize a graph $G'=(V', E')$, and $V'\leftarrow V$

Add all edges in the v-structures in G into G' (E')

```

 $E_0 \leftarrow \{e \in E, e \notin \hat{E}\} = \{e_1, e_2, \dots, e_m\}, m = |E_0|$ 
Let  $x[1..m]$  be the flags that whether  $e_i$  will be added into
 $G'$ ,  $x_i \in \{0, 1\}$ 
 $x[1..m] \leftarrow \text{BEG}(G', E_0)$ 
FOR  $i=1$  TO  $m$  DO
  IF  $x_i = 1$  THEN
     $G' \leftarrow G' \cup \{e_i\}$ 
  END IF
END FOR
RETURN  $G_M^* \leftarrow G'$ 

```

For edge selection, we adopt a max-heap, denoted as T , described as follows: *enode*: the node in T , whose left child and right child are with flags 1 and 0 respectively, describing the corresponding edge in E_0 will be selected or not; *level*: level number corresponding to current node *enode*; *parent*: the parent node of *enode*; *size*: the size of the subgraph including the selected edges from the root to *enode*; *nodes*: edges that have been selected currently; *upper*: the maximal possible size considering the nodes that have been selected and those that have not been selected.

Algorithm 4. BBG

```

Input: PDAG  $G$ , Candidate edges  $E = \{e_1, e_2, \dots, e_m\}$ 
Output:  $\{x_1, x_2, \dots, x_m\}$ ,  $x_i \in \{0, 1\}$ 
Steps:
Construct an empty max-heap  $H$ ; a binary tree  $T$ 
 $enode \leftarrow \phi$ ,  $cSize \leftarrow 0$ ,  $i \leftarrow 1$ ,  $maxSize \leftarrow 0$ ,  $upper \leftarrow 0$ ,  $cNodes \leftarrow \phi$ ,
 $rNodes \leftarrow \phi$ 
WHILE  $i \neq m+1$  DO
  IF  $\text{feasible}(G, enode.nodes \cup \{e_i\})$  THEN
     $cNodes \leftarrow enode.nodes \cup \{e_i\}$ ,  $cSize \leftarrow cSize + 1$ 
    IF  $cSize > maxSize$  THEN
       $maxSize \leftarrow cSize$ 
      InsertToHeap( $H, i+1, maxSize, cNodes, 1$ )
    END IF
  END IF
   $rNodes \leftarrow \phi$ 
  FOR  $j \leftarrow i+1$  TO  $m$  DO
     $rNodes \leftarrow rNodes \cup \{e_j\}$ 
    IF  $\text{feasible}(G, rNodes)$  THEN
       $upper \leftarrow upper + 1$ 
    END IF
  END FOR
  IF  $|enode.nodes \cup rNodes| \geq 1$  AND  $upper \geq maxSize$ 
    THEN InsertToHeap( $H, i+1, upper, cNodes - \{e_i\}, 0$ )
  END IF
   $enode \leftarrow \text{removeMaxHeap}(H)$ ,  $i \leftarrow enode.level$ 
   $enode \leftarrow enode.parent$ 
END WHILE
FOR  $i \leftarrow m$  DOWNTO 1 DO
  IF  $enode.leftchild = 1$  THEN  $x_i \leftarrow 1$  ELSE  $x_i \leftarrow 0$ 
  END IF

```



```

END FOR
RETURN{ $x_1, x_2, \dots, x_m$ }

```

The execution time of Algorithm 3 mainly depends on that of Algorithm 4, which will be $O(2^m)$ at the worst case. Actually, the average execution time of Algorithm 4 is much less than that of the worst case to a great extent due to the pruning strategies.

3.3 Fusing DAGs of BNs

We first derive the maximal equivalent subgraph G^* . Second, we consider adding possible edges in CG_2 into G^* , and then adding those in CG_1 and CG_2 . Third, based on Algorithm 2, we can obtain the result DAG by the consistent extension of G^* . The above ideas are summarized and given in Algorithm 5.

Algorithm 5. Fuse-DAG

```

Input: DAGs of the BNs  $\{G_1, G_2, \dots, G_n\}$ 
Output: DAG  $G^* = (V^*, E^*)$ 
Steps:
FOR  $i \leftarrow 2$  TO  $n$  DO
   $G_1 = (V_1, E_1) \leftarrow G_{i-1}; G_2 = (V_2, E_2) \leftarrow G_i$ 
   $CG_1 \leftarrow \text{DAG-TO-CPDAG}(G_1); CG_2 \leftarrow \text{DAG-TO-CPDAG}(G_2)$ 
   $CG^* = (CV^*, CE^*) \leftarrow CG_1 \cap CG_2 = (CV_1 \cap CV_2, CE_1 \cap CE_2)$ 
  IF the subgraphs of  $G_1$  and  $G_2$  on  $CV^*$  are not Markov
  equivalent THEN
     $G_M^* = (CV_M, CE_M) \leftarrow \text{Max-Subgraph}(CG^*)$  //Algorithm 3
  ELSE
     $G_M^* = (CV_M, CE_M) \leftarrow CG^*$  //Theorem 2
  END IF
  FOR  $i \leftarrow 1$  TO 2 DO
    FOR each edge  $e$  in  $CE_i - CE_M$  DO
      Let  $e = (x, y)$ 
      IF feasible  $(CV_M \cup \{x, y\}, CE_M \cup \{e\})$  THEN
         $V_0 \leftarrow CV_M \cup \{x, y\}, E_0 \leftarrow CE_M \cup \{e\}$ 
      END IF
    END FOR
  END FOR
   $G_0 \leftarrow (V_0, E_0), G_i \leftarrow G_0$ 
END FOR
 $G^* \leftarrow \text{PDAG-TO-DAG}(G_0)$  //Algorithm 2
RETURN  $G^*$ 

```

In the above steps, Algorithm 1 and Algorithm 2 can be invoked and executed in polynomial time. Algorithm 3 will be invoked for $O(n)$ times.

4 Conclusions and Future Work

To fuse multiple PGMs for representing and inferring uncertain knowledge, we preliminarily proposed an approach for fusing the DAGs of BNs by preserving the maximal conditional independencies. For space limitation, we just presented our ideas and methods briefly in this paper.

Accordingly, the fusion of BN's CPTs can be considered further. The underlying techniques for fusing PGMs in specific situations, e.g., time-series, group decision, etc., can be studied further. These are exactly our future work.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (No. 61063009), the Ph.D. Programs Foundation of Ministry of Education of China (No. 20105301120001), the Foundation for Key Program of Ministry of Education of China (No. 211172), and the Foundation for Key Program of Department of Education of Yunnan Province (No. 2011Z015).

References

1. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, Palo Alto (1998)
2. Li, W., Liu, W., Yue, K.: Recovering the Global Structure from Multiple Local Bayesian Networks. *International Journal of Artificial Intelligence Tools* 17(6), 1067–1088 (2008)
3. Zhang, Y., Yue, K., Yue, M., Liu, W.: An Approach for Fusing Bayesian Networks. *Journal of Information and Computational Science* 8(2), 194–201 (2011)
4. Matzkovich, I., Abramson, B.: The topological fusion of Bayes nets. In: Proc. of the 8th Conf. UAI, pp. 191–198 (1992)
5. Sagrado, J., Moral, S.: Qualitative combination of Bayesian networks. *International Journal of Intelligent Systems* 18, 237–249 (2003)
6. Richardson, M., Domingos, P.: Learning with knowledge from multiple experts. In: Proc. of the 20th Conf. ICML, pp. 624–631 (2003)
7. Nielsen, S., Parsons, S.: An application of formal argumentation: Fusing Bayesian networks in multi-agent systems. *Artificial Intelligence* 171, 754–775 (2007)
8. Verma, T., Pearl, J.: Equivalence and synthesis of causal models. In: Proc. of the 6th Conf. UAI, pp. 220–227 (1990)
9. Chickering, D.: Learning equivalence classes of Bayesian-network structures. *Journal of Machine Learning Research* (2), 445–498 (2002)
10. Andersson, S., Madigan, D., Perlman, M.: A characterization of Markov equivalence classes for acyclic digraphs. *Annals of Statistics* 25, 505–541 (1997)
11. Chickering, D.: A transformational characterization of Bayesian network structures. In: Proc. of the 11th Conf. UAI, pp. 87–98 (1995)
12. Meek, C.: Causal inference and causal explanation with background knowledge. In: Proc. of the 11th Conf. UAI, pp. 403–410 (1995)

A Hybrid Genetic Algorithm for Scheduling and Selecting a Project Portfolio

Bo Shi¹, Hong Wang², and Lu Qi¹

¹Department of Computer Science, JiNing University,
QuFu City XingTan Road No.1, ShanDong Province, P.C. China, 273155
mengboshi@139.com

²Information Science and Engineering School, ShanDong Normal University
JiNan WenHua Road No.88, ShanDong Province, P.C. China, 250014
wanghong106@163.com

Abstract. In this study, we consider the problems associated with selecting and scheduling a set of R&D projects to maximize the overall net present value. This paper proposes a zero-one integer programming model in conjunction with a genetic algorithm (GA) to overcome these problems. Taguchi Method was employed in the design of the GA parameters to increase the efficiency of the proposed method. We conclude that the proposed GA is capable of efficiently solving problems associated with the management of portfolios.

Keywords: project portfolio, Taguchi method, genetic algorithm, 0-1 integer programming.

1 Introduction

Project portfolio selection is a crucial management activity for many organizations and it is a complex decision-making process. Archer and Ghasernzadeh [1] defined project portfolio selection as the periodic activity involved in selecting the set of project, from available project proposals and projects currently underway. Generally the traditional approach of project selection includes two issues: one is to select a set of projects that meet some predetermined goals and resource constraints, and the other is to schedule this set of projects within planning horizon without violating annual budget limit. Recent works on this issue are presented by Stummer et al. [13], they developed multi-criteria decision support system (MCDSS) that first determines the set of Pareto-efficient solutions. Wang [15] proposed R&D project selection models based on linear, non-linear, dynamic, goal, and stochastic mathematical programming. Liesio [10] developed the Robust Portfolio Modeling methodology which extends Preference Programming methods into portfolio problems.

In this study, hence, a GA incorporating a new efficient experimental design method for parameter optimization using Taguchi method for the problem is proposed. Robust designs such as Taguchi method borrow many ideas from the statistical design of experiments for evaluating and implementing improvements in products, processes, or equipment. The rest of this paper is described as follows. In section 2 the problem

definition is provided and a zero-one linear programming for project selection, scheduling and balancing problem is formulated. Section 3 presents a genetic algorithm. Section 4 describe the design of experiment and evaluates the performance of the proposed GA. Section 5 summarizes the conclusions.

2 Problem Definition

2.1 Problem Statement

In this paper, our objective is to select a set of projects, organized around a set of key corporate strategies, which maximize their NPV and satisfy the strategic intent targets with balance, constrained by total and annual budget availability, and precedence relationships. The balance is measured in terms of the percent of spending directed at each strategy.

2.2 Proposed Model

We consider the situation where there are S strategic intent targets in a firm and T planning horizon periods. Suppose that there are n_j candidate projects in strategy category i . Our objective is to select a set of projects, organized around a set of key corporate strategy, which maximize their NPV assumed and satisfy the strategic intent categories with balance, constrained by total and annual budget availability, and precedence relationships. The balance is measured in terms of the percent of spending directed at each strategy. In this paper, we set the maximum and minimum fraction of the total budget that can be spent on projects contributing to achievement of strategy 1.

The decision variables are defined by:

$$Z_{Maximize} = \sum_{i=1}^S \sum_{j=1}^{n_i} \sum_{t=1}^T P_{ij} x_{ijt} \quad (1)$$

Subject to

$$\sum_{i=1}^{SC} \sum_{j=1}^{n_i} \sum_{t=1}^T w_{ij} x_{ijt} \leq TC \quad (2)$$

$$SI - LB_i \leq \sum_{j=1}^{n_i} \sum_{t=1}^T w_{ij} x_{ijt} \leq SI - UB_i \quad (3)$$

$$\sum_{i=1}^{SC} \sum_{j=1}^{n_i} \sum_{t=1}^T C_{ij,k-t+1} x_{ijt} \leq AF_k, \text{ for } k = 1..T \quad (4)$$

$$\sum_{j \in N_i} \sum_{t=1}^T tx_{ijt} + D_{ij} \leq T + 1 \quad (5)$$

$$\sum_{t=1}^T x_{ijt} \leq \sum_{t=1}^T x_{ikt}, \text{ for } j \in P_k \quad (6)$$

$$\sum_{t=1}^T tx_{ikt} + (T + 1) * (1 - \sum_{t=1}^T tx_{ikt}) - \sum_{t=1}^T tx_{ijt} \geq D_{ij} \sum_{t=1}^T tx_{ijt}, \text{ for } j \in P_k \quad (7)$$

$$\sum_{i \in HR} \sum_{j \in n_i} (w_{ij} \sum_{t=1}^T x_{ijt}) \leq PHR * (\sum_{i=1}^{SC} \sum_{j \in n_i} w_{ij} x_{ijt}) \tag{8}$$

$$\sum_{t=1}^T x_{ijt} \leq 1, \text{for } j \in EP \tag{9}$$

$$\sum_{t=1}^T x_{ijt} = 0, \text{ for } j \in EP \tag{10}$$

$$X_{ijt}=1 \text{ for } i \in IP \tag{11}$$

3 Genetic Algorithm

Before a genetic algorithm (GA) can be run, a suitable representation for the considered problem must be devised. We also require a fitness function, which assigns the project profit to each representation. During the run, parents must be selected for reproduction, and use genetic operators to generate offspring.

3.1 Direct Representation

To a direct problem representation, the problem itself is used as a chromosome. No decoding procedure is necessary. All information relevant to the problem at hand is included in the problem representation. A complete direct representation of a chromosome (i.e. project portfolio) comprises each candidate project of all strategic categories with associated X_{ij} representing whether a project is selected/non-selected and Y_{ij} denoting the starting period of the selected/non-selected project, as shown in Figure 1.

1				2			...	SC		
X_{11}	X_{12}	...	X_{1n1}	X_{21}	X_{22}	X_{2n2}	...	X_{sc1}	X_{sc2}	X_{scnsc}
Y_{11}	Y_{12}	...	Y_{1n1}	Y_{21}	Y_{22}	Y_{2n2}	...	Y_{sc1}	Y_{sc2}	Y_{scnsc}

Fig. 1. Direct representation of a chromosome

Each cell (i.e. gene) contains three elements: the upper one, n_j , represents category i ; the middle one, X_{ij} , denotes whether a candidate project j in category i is selected; the lower one, Y_{ij} , means the starting period of the selected project j in category i . Suppose there are SC strategic categories, N_1 to N_{sc} , and each category i has n_i candidate projects. If a candidate project j in category i is selected into project portfolio, then the value of X_{ij} is set as 1; otherwise, 0. If $X_{ij} = 1$, then the value of Y_{ij} is assigned randomly as a value from 1 to $T - D_{ij} + 1$; otherwise, $Y_{ij} = 0$. The fitness function is defined as the overall profit of a project portfolio and given as follows.

$$Function_{fitness} = \sum_{i=1}^s \sum_{j=1}^{n_i} P_{ij} X_{ij} \tag{12}$$

3.2 Initialization and Sorting

The initial generation of complete and consistent project portfolios can be generated as follows. For each project portfolio of init $\lfloor \frac{T}{N} \rfloor$ generation, we first randomly decide X_{ij} as 0 or 1. If X_{ij} equals 1, then the value of Y_{ij} is set randomly as from 1 to $T - D_{ij} + 1$. If the generated project portfolio is feasible when it meets the all above constraints, generate a new portfolio until the first generation is produced.

3.3 Genetic Operators

The introduction of a non-standard chromosome representation necessitates the definition of new crossover and mutation operators which are usually more complicated than traditional ones.

A. Elitism: Rather than the mechanism of reproduction of GA, the mechanism of elitism instead is applied. In the function of elitism, a certain ratio of elitist chromosomes is kept into the next generation to avoid losing larger fitness-value chromosomes.

B. Crossover: The crossover operator generates an offspring portfolio by combining features of two selected parent portfolios. The crossover point occurs at the gene point (i.e. gene point divides department) of a chromosome. That is, if a firm has SC strategic categories, there will be SC -1 possible crossover points.

C. Mutation: The mutation operator must be able to alter some information represented in the chromosome.

4 Computational Results

4.1 Generating Problem Instances

In this paper, a genetic algorithm is proposed for the considered problem. To investigate the performance of the proposed genetic algorithm, the solutions solved by the proposed GA are compared with those solved by AMPL. Each instance includes three strategic categories in which each has from 6 with increment 2, to 12 candidate projects. The required information in each problem instance includes required annual cost of each project, the risk value of each project, total budget limit of each strategic intent category, and annual budget limit of an organization. Once a problem instance is generated, it is solved by the GA algorithm. These results then can be compared with those obtained by AMPL. Time consumption and objective value difference will be compared. At first we code the proposed GA in VB with Microsoft Visual Basic 6.0, and then using AMPL to solve the generated problem instances.

4.2 Parameters' Design

Once a genetic algorithm is developed, its performance strongly depends on the parameters of GA. In this study, we select five most commonly studied GA parameters. Different parameter level causes different result even in the same problem instance.

After an extensive preliminary analysis of the algorithm, three levels for each parameter values are chosen. Selected design factors and their levels are listed in Table 1. Therefore, to decide the most suitable level of the parameters to get stable solutions is an important issue. Taguchi method then could be applied to decide these parameters.

Table 1. Design factors and level

Level	Elitism rate	Crossover rate	Mutation rate	Generation size	Generation size
Level1	79.88076164	79.83005777	79.84899446	79.79792417	79.82297967
Level2	79.81992428	79.85297404	79.83328824	79.84716448	79.8271885
Level3	79.81805289	79.83570701	79.83645611	79.87407716	79.86857064

In this study, we want to estimate the main effects of design factors (the interaction among factors is neglected). An efficient way of studying the effects of several design factors simultaneously is to plan a matrix experiment using an orthogonal array. For the inner array, we choose orthogonal array L27 (3¹³) which has 13 three-level columns and 27 rows. Each row of the inner array L27 represents a design of the process.

After confirming the orthogonal arrays, the next step is to determine the signal-to-noise ratio by the experiment data. By defining our case as a type of maximum problem, the objective function to be maximized can thus be represented by the following equations.

$$\eta = S / N = - 10 \log_{10} (MSD) = - 10 \log_{10} \left[\frac{1}{n} \sum_{i=1}^n \left(\frac{1}{y_i^2} \right) \right] \tag{13}$$

Where η is the S/N ratio for each experiment; n is the representative number of measurements and Y_i is each observation of experiment. Here Case 6 is taken as an illustration of Taguchi method. Firstly, 27 S/N ratios of Case 6 are calculated, and then the average of three levels for each parameter can be obtained, shown in Table 2.

From Table 2, A1, B2, C1, D3 and E3 are the best parameter combination to this case. The next step is input the above parameter combination to the GA system again. The output result to Case 6 under this parameter combination is 10044. In Case 6, we can obtain the fitness value of 10044 by the GA method. The following steps are the procedure of the confirmatory experiment.

Table 2. Average values of three levels for each

Parameters	Level1	Level2	Level3
Elitism rate (A)	0.2	0.4	0.6
Crossover rate(B)	0.2	0.4	0.6
Mutation rate (c)	0.2	0.4	0.6
Populationsize(D)	50	100	150
Generationsize(E)	50	100	150

Step 1. Calculate the S/N optimum

$$\eta_{opt} = \bar{\eta} + (\eta_{A1} - \bar{\eta}) + (\eta_{D3} - \bar{\eta}) + (\eta_{E3} - \bar{\eta})$$

In the above equation, $\bar{\eta} = \frac{\eta_1 + \eta_2 + \dots + \eta_{27}}{27} = \frac{2155.668649}{27} = 79.8396$

$$\therefore \eta_{opt} = 79.8808 + 79.8686 - 79.8396 - 79.8396 = 79.94425$$

Step 2. Calculate the average S/N values

In step 2, we experiment several times by using the best parameter combination to get the S/N value of the results. Here three times are performed, the average S/N value, 79.998 is obtained.

Step 3. Calculate the difference between S/N optimum and the average S/N value obtained in Step2.

$$\frac{S / N - S / N_{opt}}{S / N_{opt}} * 100 \% = \frac{79.998 - 79.94425}{79.94425} * 100 \% = 0.067 \%$$

Due to 0.067% (less than 20%), this experiment can be confirmed. In addition, ANOVA is used to calculate the contribution of each parameter of our proposed GA.

4.3 Performance Evaluation of Proposed GA

To evaluate the performance of the proposed GA, the solutions obtained by the GA are compared with those by AMPL modeling language. Therefore, the value of Difference Parameter Degree of freedom = [(AMPL value - GA value)/AMPL value] is computed. Given by the six generated problem instances, Table 3 shows that the proposed GA can find the optimal solutions as found by AMPL in small problem instances or the near-optimum solutions even for some slightly larger problem instances. Moreover, the GA solves our considered problem more efficiently than AMPL does. That is to say, the time required to find the solution found by GA is far less than by AMPL.

Table 3. Solutions found by AMPL and the proposed GA

Problem instance	GA value	AMPL value	Difference
#1	3747(0.586)*	3747(0.45)	0.00%
#2	6434(11.594)	6434(1.2)	0.00%
#3	7614(19.133)	7716(1.5)	1.32%
#4	7846(27.277)	7962(3.6)	1.46%
#5	10341(30.906)	10461(198)	3.15%
#6	10044(33.625)	10306(3507)	2.54%

5 Conclusion

The problem considered in this paper is to deal with the issue, balancing the strategic intent targets, added to a traditional project selection problem which comprises two issues. In this paper a zero-one integer linear programming model for the considered problem is proposed and a genetic algorithm (GA) is proposed to solve the problem. Furthermore, to increase the efficiency of the proposed method, GA parameter design in accordance with Taguchi Method is conducted. Some problem instances are randomly generated to evaluate the performance of the proposed method by comparing with the solutions solved by AMPL in small scale of problem instances. From the computational results, it can lead to a conclusion that the proposed GA provides an efficient solution to the problem.

References

1. Archer, N., Ghasemzadeh, F.: An integrated framework for project portfolio selection. *International Journal of Project Management* 17(4), 207–216 (1999)
2. Coffin, M.A., Taylor III, B.W.: R&D project selection and scheduling with a filtered beam search approach. *JJE Transactions* 28, 167–176 (1996a)
3. Coffin, M.A., Taylor III, B.W.: Multiple criteria R&D project selection and scheduling using fuzzy logic. *Computers and Operations Research* 23(3), 207–220 (1996b)
4. Evans, G.W., Fairbairn, R.: Selection and scheduling of advanced mISSIONs for NASA using 0-1 integer linear Programming. *Journal of the Operational Research Society* 40, 971–981 (1989)
5. Ghasemzadeh, F., Archer, N., Iyogun, P.: A zero-one model for project portfolio selection and scheduling. *Journal of the Operational Research Society* 50, 745–755 (1999)
6. Gray, C.F., Larson, E.W.: *Project Management: The Managerial Process*, 4th edn. McGraw-Hill, New York (2006)
7. Ghorbani, S., Rabbani, M.: A new multi-objective algorithm for a project selection problem. *Advances in Engineering Software* 40(1), 9–14 (2009)
8. Heidenberger, K., Stummer, C.: Research and development project selection and resource allocation: A review of quantitative modelling approaches. *International Journal of Management Reviews* 1(2), 197–224 (2003)
9. Kyparisis, G.J., Gupta, S.K., Ip, C.: Project selection with discounted returns and multiple constraints. *European Journal of Operational Research* 94, 87–96 (1996)
10. Liesio, J., Mild, P., Salo, A.: Preference programming for robust portfolio modeling and project selection. *European Journal of Operational Research* 181, 1488–1505 (2007)
11. Ozdamar, L.: A genetic algorithm approach to a general category project scheduling. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews* 29(1), 44–59 (1999)
12. Schniederjans, M., Santhanam, R.: A multi-objective constrained resource information system project selection method. *European Journal of Operational Research* 70, 244–253 (1993)
13. Stummer, C., Kiesling, E., Gutjahr, W.J.: A multicriteria decision support system for competence-driven project portfolio selection. *International Journal of Information Technology & Decision Making* 8(2), 379–401 (2009)
14. Sun, H., Ma, T.: A packing-multiple-boxes model for R&D project selection and scheduling. *Technovation* 25, 1355–1361 (2005)
15. Wang, J., Hwang, W.L.: A fuzzy set approach for R&D portfolio selection using a real option valuation model. *Omega* 35, 247–257 (2007)

Tibetan Processing Key Technology Research for Smart Mobile Phone Based on Symbian

Nyima Trashi¹, Qun Nuo¹, Yong Tso¹, Pu Dun², Gama Zhaxi¹, and Niluo Qiongda¹

¹Modern Education Technology Center, Tibet University, No.36,
Jiangsu Road, Lhasa of Tibet, China

²Teaching Affairs Office, No.36, Jiangsu Road, Lhasa of Tibet, China
{niqiongda, q_nuo, yongtso, pudun88}@163.com,
{gama, niqiong}@utibet.edu.cn

Abstract. The author has been engaging in the research of Tibetan language information processing techniques, and has presided over the development of CDMA 450M network-based Tibetan mobile phone and car phone[1],[5], as well as smart mobile phone Tibetan software package based on Windows Mobile and Android. In 2010, required by China Mobile Communications Group Co., Ltd. Tibet Branch, Symbian S60 v3 operating system based smart mobile phones was developed. Through in-depth study of Symbian operating system, the author firstly proposed Tibetan processing key technology implemented on Symbian S60 v3 based smart mobile phone.

Keywords: Symbian, Tibetan, Mobile Phone, Input/Output.

1 Introduction

With the development of mobile communication technology and wide use of mobile phones, the number of mobile phone user grows fast and the demand on Tibetan mobile phone keeps on rising. Therefore, Tibetan mobile phone software package research and development has a special significance in realizing mobile phone Tibetan processing improving the country's overall level of information, promoting socioeconomic development of minority areas.[9]

The implementation of Tibetan input/output is the core of mobile phone Tibetan processing technology.[13] This paper firstly proposes a kind of Tibetan input/output technology based on Symbian S60 v3 operating system, designs a strongly adaptable Tibetan keyboard layout, which achieves fast Tibetan input and fills the blanks of Tibetan processing in the open source operating system of smart mobile phone.

2 Symbian and Its Input/Output Technology

2.1 Symbian

The predecessor of Symbian operating system is British Psion's EPOC operating system. It is an open and standard multitasking operating system. As a very mature

operating system, Symbian has the following features: efficient memory management, multi-tasking, smooth running and switching, good openness. Currently, according to different man-machine interface, Symbian’s UI platform includes S60, S80, S90 and UIQ, etc, of which S60 platform is very popular among manufacturers and users. In the context of open source, research on Symbian operating system and its localization has become possible. At present, the main development languages supported by Symbian are Java and C++; Symbian Tibetan input/output is achieved on Symbian S60 v3.[6],[7],[8].

2.2 Symbian Input Interface

The input method in Symbian operating system is implemented by FEP (Front End Processors) interface. Located in the position between the user and the application, FEP is used by the system as a dynamic link library to implement text input through some input method. It can receive user input (key, mouse, sound, etc), and send text or symbols to application. The workflow of FEP is shown in Figure 1.

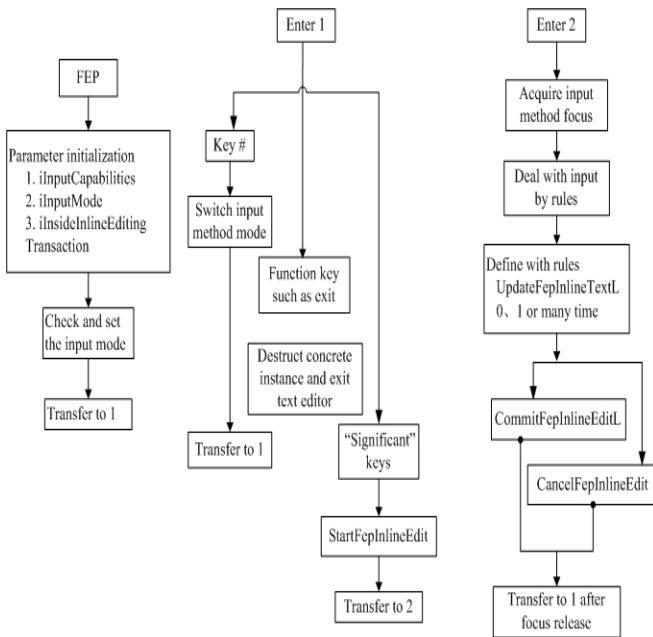


Fig. 1. FEP Workflow

3 Implementation of Symbian Tibetan Input/Output

3.1 Implementation Principle

Developers can implement the input method through inheriting basic class CCoeFep of FEP. CCoeFep has ten interface functions, of which the most important two are as follows:

```
IMPORT_C CCoeFep * NewFepL (CCoeEnv&, const TDesC&, const CCoeFepParameters);
```

```
IMPORT_C void SynchronouslyExecuteSettingsDialogL (CCoeEnv&, const TDesC&);
```

Interface NewFepL () returns to an input object pointer, this object has the highest authority for input event and can carry out input event processing before application. This interface is called in the initialization of each application. Interface SynchronouslyExecuteSettingsDialogL () returns to an input method setting dialog box.

In addition to inheriting basic class CCoeFep, Symbian Tibetan input method dynamic link library CTibetanFepPlugin still needs to achieve two important functions: constructor and HandleChangeInFocus (). The most important task of constructor is to create CTibetanControl, which is the prompt box of Tibetan input method and is used to receive user's input. When a control in the operating system has the focus, HandleChangeInFocus function is used to announce focus change to Tibetan input method. If the current control can accept text input, the input capability information includes an interface named MCoeFepAwareTextEditor. Tibetan input method calls MCoeFepAwareTextEditor interface and communicates with the current control. MCoeFepAwareTextEditor interface has four main functions: StartFepInlineEditL(), UpdateFepInlineTextL (), CommitFepInlineEditL () and CancelFepInlineEdit (). These four functions are used to control the control input status. StartFepInlineEdit is used to indicate the beginning of a text input, and then the input method calls UpdateFepInlineTextL to update the current input method status based on input status. If the input is completed, the input method calls CommitFepInlineEditL to submit input text to the current control; if the user cancels the input, CancelFepInlineEdit is called.

3.2 Input State Transition Diagram Is Shown in Figure 2

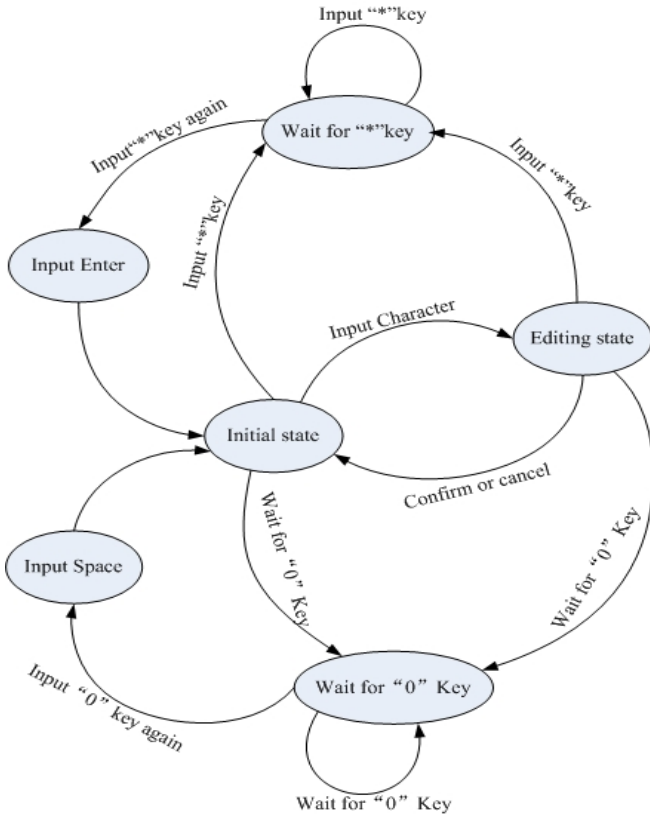


Fig. 2. Tibetan Input Method State Transition Diagram

3.3 Coding Scheme

Tibetan coding uses “Information technology-Tibetan coded character sets for information interchange, Basic set” (hereinafter referred to as the “Basic Set”) and “Information technology- Tibetan coded character set-Expansion A” (hereinafter referred to as “Expansion A”) national standards. The “Basic Set” coded character set is located in Unicode (ISO/IEC 10646 BMP) U+0F00~U+0FBF; “Expansion A” coded character set is located in Unicode (ISO/IEC 10646 BMP) 0xF300~0xF8FF, that is, the PUA area. [3],[10]

“Information technology- Tibetan coded character set-Expansion B” (hereinafter referred to as “Expansion B”) national standards collect 5702 Sanskrit Tibetan characters and all characters are encoded in a OF of GB 13000.1-1993, whose location is 0xF0000-0xF1645, and the code of each character consists of 3 bytes. Because the encoded characters in “Expansion B” are less used Sanskrit Tibetan characters, which are rarely used in handheld devices, there is no need for Tibetan input/output to support “Expansion B”. [4]

3.4 Keyboard Layout

The keyboards of most mobile phones are composed of 12 keys, in 3×4 array. To reasonably arrange Tibetan words on limited keys, Symbian Tibetan input method arranges the first 28 words of 30 Tibetan words in keys 1 to 7 every 4 of them according to Tibetan classification. According to Tibetan grammar, the 29th and 30th words have no super-fix or sub-fix but can only be spelled with vowel signs; therefore, to reduce repetition rate of input method, arrange these two words on key “8” and “9” respectively and also put the 4 vowel symbols on these two keys. Arrange the most commonly used Tibetan word delimiter in key “0” and common Tibetan symbols and numbers on key “*”; achieve “Enter” type through double-clicking and achieve “Space” character input by double “0” click. [2],[3],[5],[11],[12]

The input method designs keyboards for Tibetan and commonly used Sanskrit Tibetan characters, as shown in Figure 3. The two keyboards achieve switching by key “#”.



Fig. 3. Input Keyboard Key Position Diagram

The Tibetan shown in Figure 4 means I’m a teacher in the school. According to the Tibetan coding scheme and keyboard layout design, the corresponding keyboard input code and internal code is as follows.

Tibetan input method keyboard input code:

10,77940,175680,3190,7130,6830

Tibetan word internal code:

0F440F0B,F5D90F560F0B,F36CF5910F0B,0F51F35E0F0B,F3740F530F0B,F5A40F530F0D

ང་སློབ་གྲུའི་དགོ་ཚན་ཡིན།

Fig. 4. Example of Tibetan

3.5 Font

Symbian S60 v3 Tibetan input/output uses TrueType font, and use the third-party software to use Tibetan font file to replace system font file.

4 Instance of Input Method Application

In Symbian input/output system structure and technology environment, the author has developed and achieved Symbian S60 v3 Sunshine Tibetan Input Method. Figure 5 is an instance of Tibetan input method application in Nokia N86 smart mobile phone. The interface includes Input Method Icon, Candidate Window and Text Window. It is known from the concrete instance that the Symbian S60 v3 based Tibetan input/output technology of mobile phone is feasible and effective.

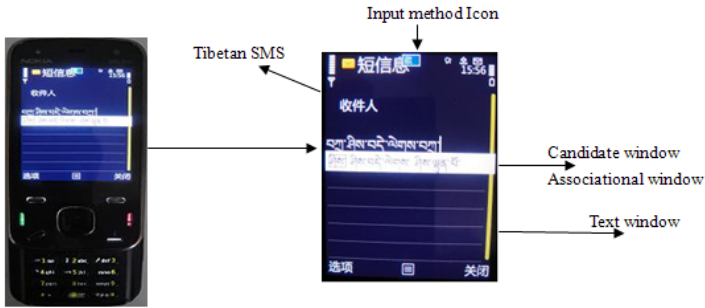


Fig. 5. Example of Input Method Application

5 Conclusion

There is significant difference between Symbian Tibetan input/output technology development and development of Windows Mobile and Android Tibetan Input/out technology. Through in-depth study on Symbian input/output system structure and its technology environment, the author first proposes a kind of Tibetan input/output technology based on Symbian S60 v3 smart mobile phones to achieve fast Tibetan input. In order to make Tibetan keyboard layout adaptable to most mobile phones, a highly adaptive Tibetan keyboard layout is designed by the own national patent, laying the basis of standard for Tibetan mobile phone keyboard layout. The technology is a kind of technological innovation, filling the blanks of related fields both at home and abroad. Symbian S60 v3 Tibetan package is the world's first one winning Symbian official certificate and digital signature.

References

1. Trashi, N., Shu, L.Z., Nuo, Q., Dun, P., Yang, C.Z., Tso, Y.: A Tibetan Mobile Phone based on CDMA System. In: Communication Systems and Networks, IASTED AsiaCSN 2007, Phuket, Thailand, April 2-4 (2007), ISBN CD:978-0-88986-658-4:24-28
2. Trashi, N., Zhi-Shu, L., Nuo, Q., Dun, P., Tso, Y., An-Long, C.: A Method of Implementing Tibetan Processing for Mobile Phone. Journal of Sichuan University(Engineering Science Edition) 41(1), 162-167 (2009)

3. Trashi, N., Zhi-Shu, L., Tso, Y., Nuo, Q., Dun, P.: Research on Key Issues of Implementing Computer Tibetan Fast Input. *Journal of University of Electronic Science and Technology of China* 38(1), 102–107 (2009)
4. Trashi, N., Zhi-Shu, L., Nuo, Q.: Computer Conversion Model of Tibetan Coding. In: *International Conference on Communications, Circuits and Systems Proceedings, Communications, Networks and signal Processing (2009)*, ISBN:978-1-4244-4887-6:562-565
5. Trashi, N.; Cai, C.: Approach of Tibetan Input, Display and Messaging Implementation on Hand-held Electronic Terminal Device. China, Patent of Invention, Patent No.: ZL, 1 0037578.5 (2005)
6. Sun, J., Liu, B., Wang, X.: Research on Symbian Operating System-Based Mobile Phone Input Method. *Computer Engineering and Application* 36 (2005)
7. Li, G., Hao, Y.: Symbian OS-based Localization Technology Research. *Micro-Computer Information (embedded and SOC)* 24(4-2) (2008)
8. Han, Z., Li, P., Zhu, Q., Qian, P.: Symbian-based Intelligent Mobile Phone Chinese Input Method Design. *Computer Application and Software* 24(9)
9. Chen, Y., Yu, S.: Tibetan Information Processing: Past, Present, and Future. *China Tibetology* 4, 97–105 (2003)
10. Rui, J., Wu, J., Sun, Y.: Study on Implementing Tibetan Operating System Based on ISO/IEC10646. *Journal of Chinese Information, Processing* 19(5), 97–104 (2005)
11. Gao, D., Gong, Y.: The Optimal Arrangement of Tibetan Character onto Keyboard. *Journal of Chinese Information, Processing* 19(6), 92–97 (2005)
12. Lu, Y.: A Study of Layout and Input Method of A General Tibetan Computer Keyboard. *Journal of Chinese Information Processing* 20(2), 78–86 (2006)
13. Lin, S., Dong, Y., Wang, S., Li, T., Pudun: A Tibetan Input Method Based on Syllable Word for Mobile Phone. In: *ICISS 2005, Second International Conference on Embedded Software and System*, Xian, China, December 2005, pp. 16–18 (2005)

PCNN Automatic Parameters Determination in Image Segmentation Based on the Analysis of Neuron Firing Time

Xiangyu Deng^{1,2} and Yide Ma¹

¹ School of Information Science & Engineering, Lanzhou University,
Lanzhou 730000, China

² Department of Electronic Information Engineering, Lanzhou Polytechnic College,
Lanzhou 730050, China

Dengxy000@126.com, ydma@lzu.edu.cn

Abstract. Pulse Coupled Neural Network (PCNN) model has been widely used in digital image processing, but its parameters determination is always under a difficult situation. This paper analyzed the firing time of the coupled linking PCNN, and indicates the difference between theoretical firing time and actual firing time when the neuron is under the influence of the neighboring neurons. By analyzing the influence of the parameters on the coupling effects of neighboring neurons, a new method of setting parameters is proposed in image segmentation. Revealing that only with the proper parameters setting, can the theoretical firing time and the actual firing time of the neuron be consistent, so that the pulse burst characteristics of PCNN can be truly realized. For Lena image segmentation, etc, the similar effects is obtained with the traditional experience parameters, and showing validity and efficiency of our proposed method.

Keywords: PCNN Model, Neuron Firing time, Neuron Firing Characteristics, Parameters Determination.

1 Introduction

The PCNN, is a neural network of pulse-coupled neurons connected with image pixels. Each image pixel is associated with a neuron of the PCNN. As shown in Fig.1, a single neuron model consists of five discrete subsystems: the coupled linking, the feeding input, the modulation, the dynamic threshold, and the firing subsystem. The state of a neuron (firing or outfiring) is dependent upon the output of firing subsystem. In PCNN model, the network performance are influenced by the linking weights matrix W , decay time constants a_E , linking weight amplification coefficients V_E and V_L , internal activity item β . So parameters determination is very important [1,2] in PCNN applications. The excitation of a neuron causes the excitation of neighboring pixels with the similar gray value, which are called synchronous pulse burst characteristics. A variety of image processing mentioned above can be realized by using this kind of characteristics.

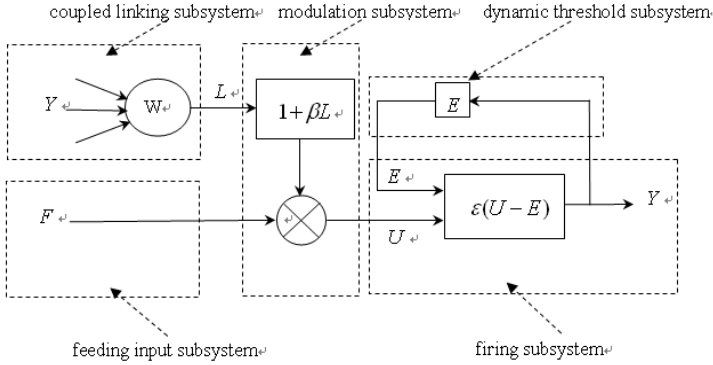


Fig. 1. Pulse Coupled Neuron Model

Recent years, automatic parameters determination of PCNN has had some research[3-6], however, most algorithms focused on the characteristics of image itself, so each method has its weak points.

2 Characteristics Analysis of the PCNN Model

2.1 The PCNN Mathematical Model

When the PCNN is applied to image processing, generally the simplified model is used, as shown in the following difference equations:

$$F_{ij}(n) = S_{ij} . \quad (1)$$

$$L_{ij}(n) = V_L \sum_{kl} W_{ijkl} Y_{kl}(n-1) . \quad (2)$$

$$U_{ij}(n) = F_{ij}(n)(1 + \beta L_{ij}(n)) . \quad (3)$$

$$Y_{ij}(n) = \mathcal{E}[U_{ij}(n) - E_{ij}(n)] . \quad (4)$$

$$E_{ij}(n) = e^{-\alpha E} E_{ij}(n-1) + V_E Y_{ij}(n-1) . \quad (5)$$

Both the coupled linking subsystem and the feeding input subsystem act on the modulation subsystem, thus the state of a neuron, which is also the output state of the firing subsystem, is dependent upon the step function $\mathcal{E}[U_{ij}(n) - E_{ij}(n)]$, and the state of this neuron also affects the neighboring neurons by the coupled linking subsystem of the neighboring neurons. When the gray value of the neighboring neurons (pixels) is close to this neuron (pixel), the neighboring pixels will be fired, which is called the capture characteristics and synchronous pulse bursts characteristics of the PCNN [7-12].

2.2 Analysis of the Neuron Firing Mechanism

From Eq.(1) to Eq.(5), gives

$$Y_{ij}(n) = \varepsilon[S_{ij} + \beta S_{ij} V_L \sum_{kl} W_{ijkl} Y_{kl}(n-1) - e^{-aE} E_{ij}(n-1) - V_E Y_{ij}(n-1)] . \quad (6)$$

Suppose that the initial value of dynamic threshold E is 0, the initial value of $Y_{ij}(n)$ is 0, and the initial value of linking subsystem is 0. Thus during the iterations the firing states of the neuron shown in Eq.(6) are discussed as follows:

1) When $n = -1$, $E_{ij}(-1) = 0$, $Y_{ij}(-1) = 0$ and $L_{ij}(-1) = 0$. So the neuron is not fired.

2) When $n = 0$, gives

$$\begin{aligned} E_{ij}(0) &= e^{-aE} E_{ij}(-1) + V_E Y_{ij}(-1) = 0 \\ L_{ij}(0) &= V_L \sum_{kl} W_{ijkl} Y_{kl}(-1) = V_L \sum_{kl} W_{ijkl} Y_{kl}(-1) = C_{ij} \\ U_{ij}(0) &= S_{ij} + \beta S_{ij} C_{ij} = S_{ij}(1 + \beta C_{ij}) \end{aligned}$$

where $Y_{kl}(-1)$ is the firing state of the neuron in the former iteration, and W_{ijkl} is the connection weights between the neuron and the neighboring. The value of C_{ij} is dependent upon the firing state of the neighboring neurons. Thus

$$Y_{ij}(0) = \varepsilon[U_{ij}(0) - E_{ij}(0)] = 1$$

Which means that the neuron is fired in the first iteration.

3) Suppose that the neuron is fired for the second time when $n = n_0$. Therefore from the Eq.(6), gives

$$e^{-(n_0-2)aE} V_E + \dots + e^{-aE} V_E Y_{ij}(n_0-2) + V_E Y_{ij}((n_0-1)) = S_{ij}(1 + \beta C_{ij}) .$$

i.e.

$$e^{-(n_0-2)aE} V_E = S_{ij}(1 + \beta C_{ij}) \Rightarrow n_0 = 1 - \frac{1}{aE} \ln \frac{S_{ij}(1 + \beta C_{ij})}{V_E e^{aE}} .$$

According to above analysis, the firing time n_m of the neuron is given in Eq.(7)

$$n_m = 1 - \frac{1}{aE} \ln \frac{S_{ij}(1 + \beta C_{ij})}{V_E (e^{aE} + e^{n_0 aE} + e^{n_1 aE} + \dots + e^{n_{m-1} aE})} \quad (m = 0, 1, 2, \dots) . \quad (7)$$

Eq.(7) shows that the firing time of the neuron is influenced by both the gray value S_{ij} and C_{ij} . In addition, for the operation of the computer has only the output at the integer time, it may cause that the different gray values have the same firing time, and the pixels with different C_{ij} are probably same in firing time, as shown in Fig.2(b). This situation will break the biological characteristics of PCNN. Fig.2 shows that the

influence of C_{ij} on the firing time have to do with the parameters a_E and V_E . When a_E tends to be smaller, the theoretical firing time and the actual firing time decided by C_{ij} are to be more concordant. In this way of setting parameters, the mathematical characteristics of PCNN and the actual operation of the computer will coincide, so that the pulse burst characteristics of PCNN will be truly realized.

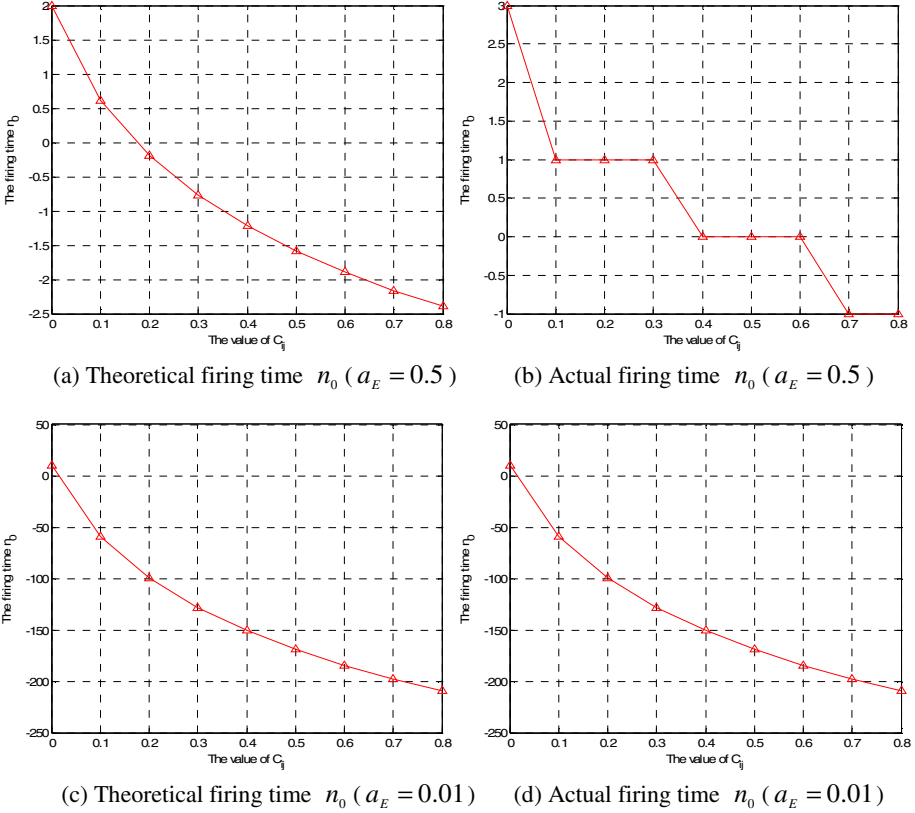


Fig. 2. The firing time n_0 with the different C_{ij} ($S_{ij} = 1, \beta = 1, V_L = 1, V_E = 1$)

3 Parameter Determination of the PCNN

In the above analysis, we supposed the gray value $S_{ij} = 1$, it stated that when $a_E \leq 0.01$, the theoretical firing time and the practical firing time can show a tendency to coincide. In addition, according to Eq.(7), the second firing time of the neuron can be gained

$$n_2 = 1 - \frac{1}{a_E} \ln \frac{S_{ij}(1 + \beta C_{ij})}{V_E e^{a_E}} . \quad (8)$$

From Eq.(8), we can find that the bigger the gray value, the smaller the firing time when the parameters of PCNN keep constant. In order to raise the pulse burst efficiency of PCNN, the maximum gray value of the image should be fired at the second iteration, gives

$$n_2 = 1 - \frac{1}{a_E} \ln \frac{S_{\max}(1 + \beta C_{ij})}{V_E e^{a_E}} \leq 2 .$$

i.e.

$$V_E \leq S_{\max}(1 + \beta C_{ij}) . \quad (9)$$

And by the image segmentation experimental, we found that when the value of V_E and normalization maximum gray value of the image are equal to each other, the preferable segmentation result can be achieved, and in this situation V_E also meets Eq.(9). By the above analysis, we obtain the way of setting parameters a_E and V_E , so the algorithm flow of PCNN image segmentation can be given:

- 1) Calculate the maximal gray value S_{\max} of the image;
- 2) Let $\beta = 1$, $V_L = 1$, $W = [0.1 \ 0.1 \ 0.1; 0.1 \ 0.1 \ 0.1; 0.1 \ 0.1 \ 0.1]$;
- 3) Let the theoretical firing time be n_{0t} and the actual firing time be n_{0a} .

when a_E decreasing from 0.5 in step of 0.01, make C_{ij} increasing from 0 to 0.8

in step of 0.1, when $\sum_{C_{ij}=0}^{0.8} \left| \frac{n_{0t} - n_{0a}}{n_{0a}} \right| \leq \varepsilon$, a_E can be got;

- 4) Let $V_E = S_{\max}$;
- 5) Start the coupled linking PCNN, and stop it when the cross-entropy of an iteration output is greater than the previous iteration.

4 Experimental Simulation and Analysis

We use a great deal of images to test the image segmentation algorithm and compared with the traditional experiment parameters. Cameraman image, for example, let $\varepsilon = 0.5$, the experimental results are shown in Fig.3, and PCNN is stopped in the fifth iteration. Fig.4 shows segmentation result comparisons of our proposed method with traditional parameters, it shows that our algorithm has obtained the same results and faster speed for only computing two parameters.

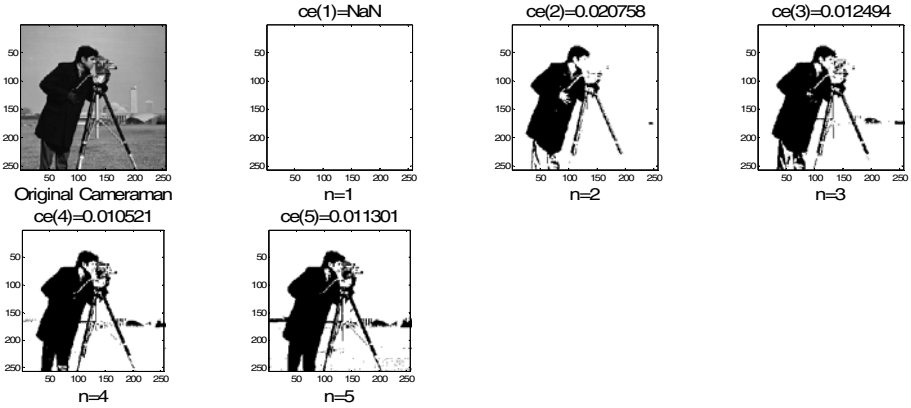


Fig. 3. Image segmentation of this paper ($ce(n)$ is cross-entropy of the n^{th} iteration)

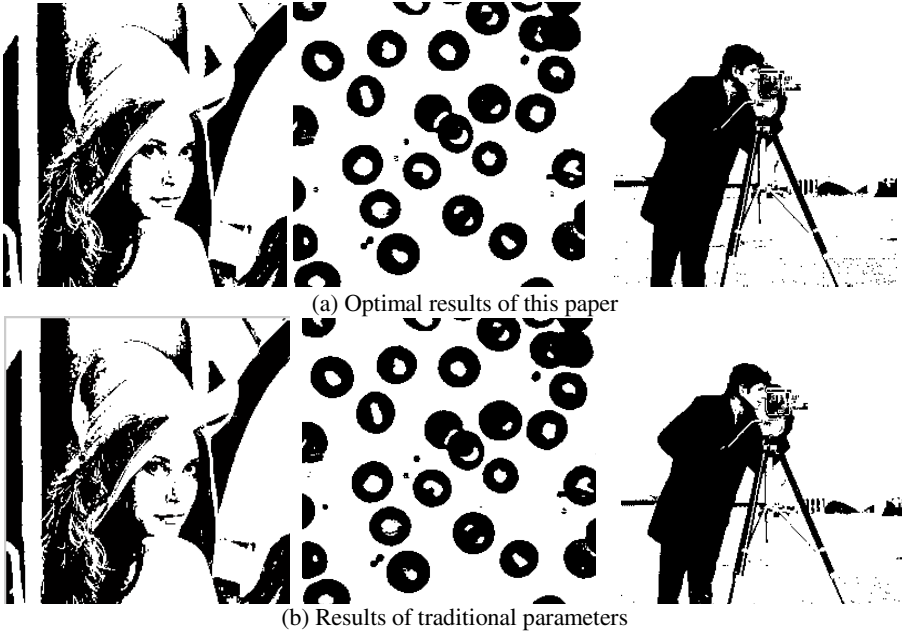


Fig. 4. Segmentation result comparisons of our proposed method with traditional parameters

5 Conclusions

Image processing of PCNN is based on the characteristics that the different gray pixels have different firing time. This paper has analyzed the firing time of the coupled linking PCNN, and indicates the difference between theoretical firing time n_{0t} and actual firing time n_{0a} when the neuron is under the influence of the neighboring

neurons. This paper also analyzes that how the parameter influences on the coupling effects of neighboring neurons. And discovering that only with the proper parameter setting, can the theoretical firing time and the actual firing time of PCNN be similar, so that the truly realization of the pulse burst characteristics of PCNN will be assured. By the experiment simulation, our algorithm obtains the similar effects compared with the traditional experiment parameters, and it also has faster speed. But this algorithm has normalized other parameters, and these parameters influences on the PCNN characteristics will be researched in the further.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grant No.60872109), the Natural Science Foundation of Gansu, China (Grant No.0710RJZA015), and the Education Department Graduate Tutor Project of Gansu, China (Grant No.1014-02).

References

1. Ma, Y., Li, L., Zhan, K.: The PCNN and Digital Image Processing. Science Publishing House, Beijing (2008)
2. Eckhorn, R., Frien, A., Bauer, R., et al.: High Frequency Oscillations in Primary Visual Cortex of Awake Monkey. *Neuro. Rep.*, 243-246 (1993)
3. Wei, S., Hong, Q., Hou, M.: Automatic Image Segmentation Based on PCNN with Adaptive Threshold Time Constant. *Neurocomputing*, 1485–1491 (2011)
4. Kuntimad, G., Ranganath, H.S.: Perfect Image Segmentation using Pulse-Coupled Neural Networks. *IEEE Transactions on Neural Networks*, 591–598 (1999)
5. Li, M., Cai, W., Tan, Z.: Adaptive Parameters Determination Method of Pulse Coupled Neural Network Based on Water Valley Area. In: King, I., Wang, J., Chan, L.-W., Wang, D. (eds.) *ICONIP 2006. LNCS*, vol. 4233, pp. 713–720. Springer, Heidelberg (2006)
6. Ma, Y.D., Liu, Q., Qian, Z.B.: Automated Image Segmentation Using Improved PCNN Model Based on Cross-Entropy. In: *Proc. Int. Symp. Intell. Multimedia Video Speech*, pp. 743–746 (2004)
7. Yu, J., Chen, H., Wang, W., Li, J.: Parameter Determination of Pulse Coupled Neural Network in Image Processing. *Acta Electronica Sinica*, 81–85 (2008) (in Chinese)
8. Bi, Y., Qiu, T.: An Adaptive Image Segmentation Method Based on A Simplified PCNN. *Acta Electronica Sinica*, 647–650 (2005) (in Chinese)
9. John, J.L., Ritter, D.: Observation of Periodic Waves in a Pulse-coupled Neural Network. *Opt. Lett.*, 1253–1255 (1993)
10. Ma, Y., Su, M., Chen, R.: Image Binarization Based on PCNN and Corresponding Segmentation Evaluation Method. *Journal of South China University of Technology (Natural Science Edition)*, 49–53 (2009)
11. Eckhorn, R.H., Reitboeck, J., Arndt, M., et al.: Feature Linking Via Synchronization among Distributed Assemblies. *Simulation of Results from Cat Cortex. Neural Comput.*, 293–307 (1990)
12. Johnson, J.L., Padgett, M.L.: PCNN Models and Applications. *IEEE Trans.on Neural Networks*, 480–498 (1999)

Messy Genetic Algorithm for the Optimum Solution Search of the HTN Planning

Jiangfeng Luo, Cheng Zhu, and Weiming Zhang

Science and Technology on Information Systems Engineering Laboratory,
National University of Defense and Technology, Changsha, P.R. China
nudtluojiangfeng@gmail.com, {chengzhu, weimingzhang}@nudt.edu.cn

Abstract. The classic algorithms of Hierarchical Task Network (HTN) planning focus on searching valid solutions with knowledge learning or heuristic mechanisms. However, there is little work on the planning optimization, especially on handling the problem in the situation that there is little heuristic knowledge to guide the optimum solution search process while many non-optimum solutions exist, or there is significance interactions between the abstract intermediate goals as which are not independent. This paper put forward a messy genetic algorithm (MGA) to solve the optimum solution searching problem of HTN planning in the above situations. Length-variant chromosome is introduced to represents the possible planning solution in form of decomposition tree with dynamic node numbers. Simulation results indicate that the MGA can locate the optimum solution among the huge search space with about 3×2^{14} possible solutions within 6 seconds.

Keywords: Hierarchical Task Network, Messy Genetic Algorithm, Optimum Solution Search.

1 Introduction

Hierarchical Task Network (HTN) [1] is a hierarchical abstraction model, which treats planning as the mission-performing process. It decomposes the high-level goal and the intermediate compound tasks (abstract level) into the primitive tasks that can be executed directly (concrete level). The execution of a primitive task is called an “action”, and the planning process is to find a course of actions (COAs) with total or partial order constraints that can achieve the high-level goal mission. Based on this ideal, the planner as SHOP2 [2] are developed. However, the HTN based planner is baffled by the follows problems:

(1) The compound tasks in the abstract level are always dependent with each other in practical missions. If we want to search the (optimum) valid solution for a HTN problem, we always baffled by the too much backtracking operation.

(2) As the number of the possible solutions for the same HTN planning problem increases exponentially with the number of decomposition methods of each compound task node, the optimum solution search of the HTN planning problem becomes a NP-complete problem [8].

2 Related Works

In HTN planner, only the methods whose preconditions are satisfied can be selected during the decomposition process of the high-level goal missions. Therefore, the constraint validation for method is a key problem for HTN planning [3]. In order to search the desired solution, knowledge learning and Heuristic method are introduced into the HTN planners [4, 5]. However, because those algorithms always offer preferable decomposition strategy based on current planning states, a large number of backtracking operations are needed. SONG et al [6] adopt a segmented backtracking method, which can builds a number of plans quickly by utilizing pieces of action-sequence produced during search process. This algorithm can provide powerful support to optimum plan generation and plan evaluation in ordered task decomposition (OTD)-based [7] HTN planning. As to reduce the search space, Reiko Tsuneto et al [9] preserve all possible decomposition strategies in an AND/OR graph, and prune the invalid branches with domain knowledge to dwindle the search space. The classic planning system SHOP2 [2] uses Branch-and-Bound optimization algorithm to guide the decomposition of the HTN to the preferable solution. It is a heuristic search algorithm requiring the domain knowledge for Branch-and-Bound. Recently, Chad Hogg et al [10] generate high quality plans through learning HTN methods.

This paper addresses the optimum solution search problem in HTN planning when there is little domain knowledge to guide the search process, and there is significance interaction between the abstract compound tasks. In section 3, we formalize the problem and build the basic model. MGA is described in section 4, in which length-variant chromosome is introduced to represents the possible planning solution in form of decomposition tree with dynamic node numbers. Section 5 is the simulation experiments, the results of which illustrates that the MGA is a feasible approach for the optimum solution search of HTN planning. Section 6 is the conclusions.

3 Problem Formation and the Basic Model

Definition 1. For the high-level goal mission G , by using the HTN methods set M to decompose it completely with constraint satisfied, we define the completely decomposed multi-level task network as the ordinary decomposition tree. All the ordinary decomposition trees of G construct the set $Tr(G)$.

The ordinary decomposition tree is the general solution of the planning, which can be denoted as:

$$tr = \{root(tr), nodeset(tr), methodset(tr), constrset(tr)\}$$

in which, $root(tr)$ is the root node of the tree while $root(tr) = G$. $nodeset(tr)$ is the node set containing all the task node of tr with the non-leaf nodes representing the compound tasks and the leaf nodes representing the primitive tasks. $methodset(tr)$ represents the methods set which contain all the selected methods during the decomposition process. $constr(tr)$ is the aggregation that contains all the constraints among the node of $nodeset(tr)$ and the preconditions of the whole

selected methods contained in $methodset(tr)$. Evidently, each level of the decomposition tree is a task network.

Definition 2. If $tr \in Tr$ is an ordinary decomposition tree of G , and the leaf task nodes of tr compose the optimum COA that can be executed to achieve the goal G , we define tr as the optimum decomposition tree. All the optimum decomposition trees of G construct the set $Tr_o(G)$.

Definition 3. Suppose there are two ordinary decomposition tree tr_i and tr_j of G with two task compound nodes $t_{ui} \in nodeset(tr_i)$ and $t_{uj} \in nodeset(tr_j)$. Evidently, t_{ui} and t_{uj} can be completely decomposed into two subtrees $subtr_i \in tr_i$ and $subtr_j \in tr_j$ with $root(subtr_i) = t_{ui}$, $root(subtr_j) = t_{uj}$. We call $subtr_i$ and $subtr_j$ are the equivalence subtrees if the follows condition satisfied:

- (1) t_{ui} and t_{uj} represent the same task;
- (2) $methodset(subtr_i) \neq methodset(subtr_j)$;
- (3) all the constraints contained in $constrset(subtr_i)$ is consistent with $constrset(tr_j)$ and all the constraints contained in $constrset(subtr_j)$ is consistent with $constrset(tr_i)$.

Especially, if $methodset(subtr_i) = methodset(subtr_j)$ and the other two conditions are satisfied, $subtr_i$ and $subtr_j$ are the same trees. The existence of the equivalence subtrees means that a compound task can be finished through different manners.

Theorem 1. Exchanging two equivalence subtrees of two different ordinary decomposition trees will get two new ordinary decomposition trees:

Proof: Because the resultant trees are completely decomposed and all the constraint conditions in each resultant tree are conflicts-free, theorem 1 is always true.

4 MGA For Optimum Solution Search of HTN

During the decomposition process of G , many compound task nodes may have multiple available decomposition methods, which makes the number of nodes and layers of the decomposition trees change dynamically. If we use the chromosomes to encode the decomposition tree, the length of the chromosome should be variable.

Designation of Initial Population. Suppose that $u \in U$ is a task node of the ordinary decomposition tree tr_i and t_u is the corresponding task, we have:

(1) if t_u is a compound task, the decomposition method set M_{t_u} of t_u is: $M_{t_u} = \{m_i | task(m_i) = t_u\}$. As encoding the decomposition tree tr_i by a length-variant chromosome, we treat t_u as gene locus ID and the actually selected decomposition method $m_i \in M_{t_u}$ as the genetic information. t_u and m_i compose a gene unit.

(2) if t_u is a primitive task with $M_{t_u} = \emptyset$, function $f(t_u)$ is defined as the utility function, which evaluate the execution cost or effect of primitive task t_u . t_u and $f(t_u)$ also compose a gene unit.

We define the chromosome of the ordinary decomposition tree as follow:

$$Chromosome(i) = \{(t_{ui_k}, gene_{t_{ui_k}}) | t_{ui_k} \in nodeset(tr_i), k \in [1, N_i], 1 \leq i_k \leq N\}$$

while N_i is the length of $Chromosome(i)$, which is equal to the number of nodes in tr_i . $(t_{ui_k}, gene_{t_{ui_k}})$ is a gene unit with t_{ui_k} is the gene locus ID and $gene_{t_{ui_k}}$ is its genetic information. We have:

$$gene_{t_{ui_k}} = \begin{cases} m_{t_{ui_k}} (m_{t_{ui_k}} \in M_{t_{ui_k}}) & \text{if } t_{ui_k} \text{ is a compound task} \\ f(t_{ui_k}) & \text{if } t_{ui_k} \text{ is a primitive task} \end{cases}$$

The gene unit ranked in the chromosome is based on the tree's root-first traverse algorithm. Then, the whole nodes of a subtree are ranked in the near offside of the subtree's root node, through which all the nodes of a subtree can compose a continuous gene segment.

The MGA starts with a randomly selected chromosome returned by the Abstract-HTN algorithm as the initial population that encodes a set of possible solution.

Crossover Operator. In MGA, crossover operator is composed of cutting operator and splicing operator [11]. As to ensure that the new chromosomes generated by cutting and splicing operation still represent the ordinary decomposition trees of HTN, we should strictly control the positions of the cutting and splicing operation. As the nodes of a subtree compose a continuous gene segment in the chromosome, exchanging the subtrees of two ordinary decomposition trees can be handled by the replacing of two gene segments with each other. Therefore, during the MGA process, the cutting operation is searching the equivalence subtrees of two parents' ordinary decomposition trees, and then, cutting out the gene segment of the equivalence subtrees in each chromosome. Splicing operator is replacing the two gene segments with each other. Based on Theorem 1 we can infer that the offspring chromosomes are also represent two ordinary decomposition trees of the HTN planning problem.

Mutation Operator. Mutation operator stochastically selects a non-leaf node in the chromosome and reconstructs the subtree whose root is the selected node by Abstract-HTN algorithm. Because the order of the task nodes stored in its chromosome

is root-first traverse sequence, the mutation operation makes the gene segment where the subtree stored mutate stochastically. According to the properties of HTN, the higher the non-leaf node's layer (in the ordinary decomposition tree) is, the more offspring nodes the non-leaf node has, and the more remarkable gene mutation takes place. As the mutation process take the constraint validation into account, the mutated chromosome still represents the ordinary decomposition tree of the HTN.

Fitness Function. Fitness function is used to measure whether the possible solutions of the planning is good or not. Because a possible solution of the HTN planning is a COA, which is composed by the primitive tasks (leaf tasks) of an ordinary decomposition tree, the fitness function of the possible solution can be defined as:

$$F(COA) = \sum f(t_u)$$

in which, t_u is primitive task, $f(t_u)$ is the measurement of the task's executing result as defined in section 4.1, such as the consumption of the resource or execution effect and so on. $F(COA)$ is the comprehensive evaluation of all the primitive tasks.

5 Simulation Experiments

We construct a HTN with the max layer number $N=6$. The number of nodes in each layer is no more than 2^{N-1} . We define that the resource consumption of one primitive task is one unit. Primitive task t_{46} has resource conflict with t_{56} , and $t_{6,16}$ have resource conflict with $t_{6,27}$, they can't exist in the same COA in couple. Evidently, there is little heuristic information can be used to guide the decomposition process. Now, we use the MGA to search the solution with the least resource consumption. During the algorithm, the selection operator is that the optimum parent individual is reserved to the offspring and the worst is eliminated directly. The size of initial population for MGA is 20, Crossover probability is 0.8, mutation probability is 0.2, and genetic generations are 300, the fitness function is finding a solution with least resource consumption. The search processes is showed in Fig1. The computer is Intel(R) Core(TM) 2 CPU 6320 @ 1.86GHz. The searching process last not more than 6 seconds.

The MGA return the optimum solution with the least resource consumption of 20 units. It shows that the MGA can locate the optimum solution among the huge search space effectively when some of the solutions of the abstract compound tasks are not independent.

6 Conclusions

MGA is a feasible approach for optimum solution search problem in HTN planning. For practical planning, we need to search the optimum COA that takes the temporal, resource and the execute effect into account. It is difficult to locate the optimum solution if there is little knowledge for heuristic algorithms to guide the optimum solution search process. The MGA discussed in this paper can give some contribution

to this problem. Furthermore, as some subtrees contain the same topology structure, the encoding method and crossover, mutation operations discussed above may be used for searching the smallest tree in the network. We should do further research about this in future.

References

- [1] Erol, K., Hendler, J., Nau, D.: HTN Planning: Complexity and Expressivity. To appear in Proc. AAAI (1994)
- [2] Nau, D., Au, T.-C., Ilghami, O., Kuter, U., William Murdock, J., Wu, D., Yaman, F.: SHOP2: An HTN Planning System. *Journal of Artificial Intelligence Research* 20, 379–404 (2003)
- [3] Surynek, P., Barták, R.: Encoding HTN Planning as a Dynamic CSP. In: van Beek, P. (ed.) CP 2005. LNCS, vol. 3709, pp. 868–868. Springer, Heidelberg (2005)
- [4] Ilghami, O., Nau, D.S., Muñoz-Avila, H., Aha, D.W.: Learning preconditions for planning from plan traces and HTN structure. *Computational Intelligence* 21(4), 143–388 (2005)
- [5] Ilghami, O., Nau, D.S., Muñoz-Avila, H.: Learning to do htn planning. In: Proceedings of the Sixteenth International Conference on Automated Planning and Scheduling, pp. 390–393 (2006)
- [6] Song, J.-G., Cha, J.-Z., Lu, Y.-P.: Multiple Plans Oriented OTD Algorithm for HTN Planning. *Journal Of Beijing Jiaotong University* 33(4) (2009)
- [7] Nau, D., Aha, D.W., Muñoz-Avila: Ordered task decomposition. In: Gil, Y., Myers, K. (eds.) Representational Issues for Real-World Planning Systems Papers from the AAAI Workshop (Technical Report WS-0018), AAAI Press, Menlo Park (2000)
- [8] Tsuneto, R., Nau, D., Hendler, J.: Plan-Refinement Strategies and Search-Space Size. In: Proceedings of the European Conference on Planning (ECP), pp. 414–426 (1997)
- [9] Tate, A.: Generating project network. In: Proceeding of the International Conference on Artificial Intelligence (IJCAI), pp. 215–218 (1975)
- [10] Hogg, C., Kuter, U., Muñoz-Avila, H.: Learning Methods to Generate Good Plans: Integrating HTN Learning and Reinforcement Learning. In: Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (2010)
- [11] Wang, X.-P., Cao, L.-M.: Genetic Algorithm: Theorem, Application and Software Practice. XI'AN JiaoTong University Press (2002)

Research on Rule-Based Reasoning Methods Oriented on Information Resource Ontology

Gang Liu^{1,a}, Lifu Feng^{1,b}, Ying Liu^{1,c}, and Zheng Wang²

¹ College of Computer Science and Technology, Harbin Engineering University, Harbin

^a liugang@hrbeu.edu.cn, ^b s310060003@hrbeu.edu.cn,

^c s310060112@hrbeu.edu.cn

² Software Engineering, University of Sydney

Abstract. This paper studies semantic retrieval method based on rule-based reasoning under support of ontology. In the support of ontology knowledge base, this system faces ontology structure-based metadata, according to related concepts and reasoning rules, looks for implication relationship and digs out implicit information. We extract reasoning rules from semantic relationship of audit information, and present Smallest Connected Subgraph Generation Algorithm, Connected Componets Generation Algorithm and Connected Componets Matching Rules Algorithm; these three algorithms are in solving problems in the process of reasoning, implement the reasoning function. Through case analysis, we verify the application probability of these algorithms in audit information resource retrieval.

Keywords: information resource retrieval, ontology, rule-based reasoning.

1 Introduction

In contrast to keyword matching-based traditional information retrieval, intelligent retrieval introduces semantic processing in the retrieval process [1], [2]. Semantic, which computers can handle, is the symbolic logic relationship between objects, these logic symbols represent the concepts of natural language [3]. After abstracting the concepts, semantic retrieval gets the true intent of the users, uses semantic relationship between concepts and reasoning, semantic information retrieval digs out hidden knowledge in information repository.

In the past, auditors use various types of audit information resources through manual search. To help auditors using variety types of related resources easily and quickly, this paper studies semantic retrieval methods based on rule-based reasoning under support of ontology and their implementations in the audit guidance systems.

2 Background and Description of the Problem

Throughout the system, ontology is information organization framework of the system. Analyze information resources, annotate domain resources according to the ontology and the formed the meta-data formation is to be stored in the metabase. Semantic reasoning, retrieval system calls inference engine, uses ontology concepts,

instances (metadata), reasons according to the relationship of concepts and inference rules, generate the results and add them to the ontology library [4]. The system search results in the ontology library for the user's query request. As ontology the repository is metadata which is the result of reasoning, matching results will contain implicit conclusion in accord with the rules of logical reasoning, so that reveals the inner profound relationship of things, implements intelligent retrieval [5].

3 Reasoning Algorithms in the Semantic Search

3.1 Initial Query of the Ontology and Ontology Smallest Connected Subgraph Generation

Definition 1: ontology O is looked as a directed graph $\langle V, E \rangle$, if for any two nodes x and y , there is an edge of x point to y , this is expressed as $\text{conn}(x, y)$;

Definition 2: ontology O is looked as an undirected graph $\langle V, E \rangle$, if for any two nodes x and y , there is an edge of x point to y or an edge of y point to x , this is expressed as $\text{conn}(x, y)$;

Definition 3: In ontology O , if there is an edge starts from x to y , it is expressed as $\text{edge}(x, y)$.

Algorithm 1.1: Smallest Connected Subgraph Generation

Input: ontology O ; initial query node set S ;

Output: the smallest connected subgraph G in O which contains all nodes in S

```

Graph G
while S != NULL do
  for each node x in S do
    remove x from initial query node set S
  procedure a(x)
  begin
    for each y that conn(x,y) then
      if y is in initial query node set S then
        add edge(x,y) into Graph G
        remove y from initial query node set S
      a(y) // recursion
    if edge(x,y) in the path is not located between two
    nodes in S then
      remove edge(x,y) from Graph G
  end

```

According to algorithm 1.1, we get Theorem 1.

Theorem 1: In the directed ontology O , if there are x, y, z , where x, y is in the initial query node set S , if $\text{conn}(x, z)$, $\text{conn}(z, y)$ and $\neg \text{conn}(x, y)$, then $\text{edge}(x, z)$ and $\text{edge}(z, y)$ are in the smallest connected subgraph.

According to algorithm 1.1, we will get the graph which contains all nodes in the initial query node set S (subgraph of the original ontology O), but in some special

cases, the resulting graph is not a connected graph. Aim at this limitation, we propose algorithm 1.2, which is similar to algorithm 1.1, just considered the directed connection $\text{conn}(x, y)$ as an undirected connection to calculate the smallest connected subgraph.

3.2 The Establishment of Rules

This paper studies ontology-support semantic retrieval, the establishment of rules mainly depends on all relevant relations between resources. Rules are formed of first order predicate logic that, in the retrieval process they are integrated into the ontology inference engine to support semantic reasoning.

Ontology-support reasoning retrieval, the establishment of rules based on semantic relations of specific domain ontology and description of the external properties in the resource, more complex inference rules, more in-depth semantic logic level search will be, retrieve accuracy and breadth will be improved.

3.3 Rule-Based Reasoning

Based on the smallest connected subgraph, we achieve rule-based reasoning. Start from each node of the smallest connected subgraph, travel the smallest connected subgraph, generating a collection of connected components (algorithm 2.1 and algorithm 2.2), then traverse each connected component, and match the rules in the rule base (see Algorithm 3), obtain reasoning results if matching successfully.

Algorithm 2.1: Connected Components Generation Algorithm

Input: smallest connected subgraph G; initial query node set S;

Output: a collection of connected components T

```

connected component K
for each node x in initial query node set S do
  procedure a(x,K)
  begin
    if conn(x,y) then
      add edge(x,y) into connected component K;
      a(y,K); // recursion
      remove edge(x,y) from connected component K;
    else
      add connected component K into a collection of
      connected components T;
  end
end

```

According to algorithm 2.1, we will get all the connected components of a given smallest connected graph, but in some special cases, the collection of connected components generated by algorithm 2.1 may be empty. Aim at this limitation, we propose the generated connected components auxiliary algorithm 2.2, when the connected components set we obtain in accordance with algorithm 2.1 is empty, we generate a collection of connected components set by reversely traversing the smallest connected subgraph.

Definition 4: In a rule base R, the antecedent of a rule r is expressed as “before”, that consequent is expressed as “after”.

Algorithm 3: Connected Components Matching Rules Algorithm

Input: connected component K; rule base R; initial query node set S;

Output: reasoning result Q

```

Result Q
temp reasoning resultset T
x is the head node of K
y is the next node of x in K
procedure a(x, y)
begin
  for each r in R do
    if edge(x, y) belongs to r.before and in r.before
      this term has not been satisfied
    then
      the number of terms in r.before minus 1;
      if the number of terms in r.before is 0
      then
        add r.after into Q;
        add r.after into T;
  a(y,y->next); // recursion
while T! = NULL do
  for each edge(x,y) in T do
    for each r in R do
      if edge(x,y) belongs to r.before and in
        r.before this term has not been satisfied
      then
        the number of terms in r.before minus 1;
        if the number of terms in r.before is 0
        then
          add r.after into Q;
          add r.after into T;
end

```

4 Application of Rule-Based Reasoning in a Retrieval System**4.1 Time Complexity Analysis**

For algorithm 1.1 and algorithm 1.2, suppose size of initial query node set S is n, and in ontology the maximum number of the nodes connected to a node in S is k, in the worst case, the time complexity of algorithm 1.1 and algorithm 1.2 is $O(kn)$.

For algorithm 2.1 and algorithm 2.2, suppose size of initial query node set S is n, each node in S will be in a connected component, in the worst case, the time complexity of algorithm 2.1 and algorithm 2.2 is $O(n)$.

For algorithm 3, suppose the number of connected components is n, the maximum number of nodes in a connected component is k, the number of rules is m, the maximum number of antecedents of a rule is t, matching rules by a connected component, the time complexity is $O(kmnt)$, the results generated by matching a rule also can be prerequisites used to match the antecedent of other rules, the time complexity is $O(m^2tn)$, so the total time complexity is $O((k + m)mnt)$.

4.2 Experimental Results

In order to validate this study, we established a retrieval experiment system, on the basis of building audit policy ontology, test these algorithms. An ontology structure fragment is shown in Figure 1.

The system performs the reasoning task, and gets the reasoning results. Suppose the user wants to know the description of “payment sum” in our national medical policy, the reasoning process and results is shown in Figure 2.

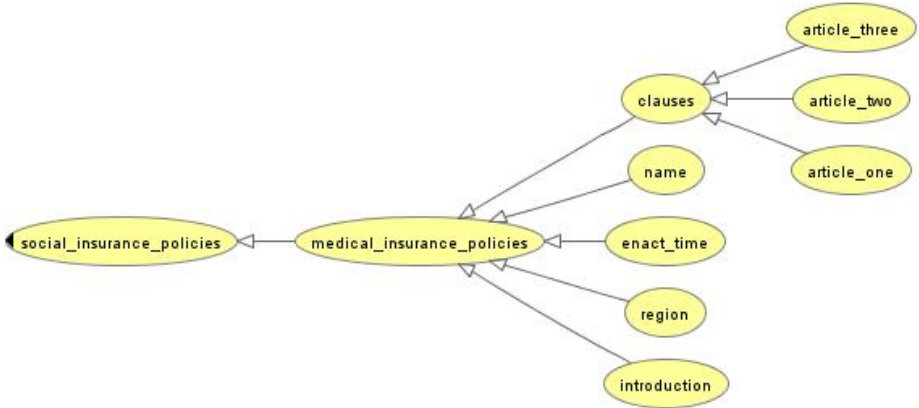


Fig. 1. An audit ontology structure fragment

W1:Payment sum W2:Payment base W3:Payment proportion
 P1:the determination of the establishment of basic medical insurance
 A1:Article 3 A1:Article 6 A1:Article 8

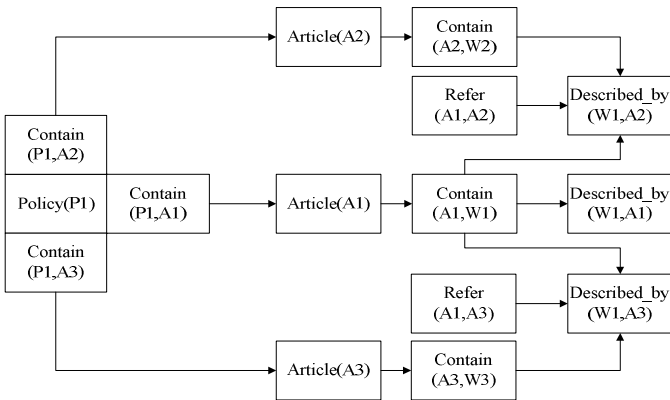


Fig. 2. Reasoning result example

5 Conclusion

This paper mainly discusses the rule-based reasoning methods and their implementation in intelligent retrieval system. The results show that, based on resource semantic relationship definition-oriented rules set, start from the existing facts to dig out the implicit conclusion in the existing facts, the algorithms are easy to implement, and are suitable to reasoning retrieval problems in audit information guidance. We will further focus on reasoning performance optimization of algorithms and improving the results.

Acknowledgment. This work is sponsored by the National Natural Science Foundation of China under grant number 60873038, the Fundamental Research Funds for the Central Universities of China under grant number HEUCF100603 and the National Science & Technology Pillar Program under grant number 2009BAH42B02.

References

1. Pessoa, R.M., Calvi, C.Z., Filho, J.G.P.: Semantic Context Reasoning Using Ontology Based Models. In: Proceedings of the 13th Open European Summer School and IFIP TC6.6 Conference on Dependable and Adaptable Networks and Services (July 2007)
2. Baggi, M., Falaschi, M., Ballis, D.: XML Semantic Filtering via Ontology Reasoning. In: The Third International Conference on Internet and Web Applications and Services (2008)
3. Drabent, W., Wilk, A.: Extending XML Query Language Xcerpt by Ontology Queries. In: IEEE/WIC/ACM International Conference on Web Intelligence (2007)
4. Bhatt, M., Rahayu, W., Soni, S.P., Wouters, C.: Ontology driven semantic profiling and retrieval in medical information systems. *Web Semantics: Science, Services and Agents on the World Wide Web* 7(4), 317–331 (2009)
5. Lourenço, A., Carreira, R., Glez-Peña, D., Méndez, J.R., Carneiro, S., Rocha, L.M., Díaz, F., Ferreira, E.C., Rocha, I., Fdez-Riverola, F., Rocha, M.: BioDR: Semantic indexing networks for biomedical document retrieval. *Expert Systems with Applications* 37(4), 3444–3453 (2010)

Generating Descriptions of Incomplete City-Traffic States with Agents^{*}

Grzegorz Popek^{1,2}, Ryszard Kowalczyk¹, and Radoslaw P. Katarzyniak²

¹ Faculty of ICT, Swinburne University of Technology, Melbourne, Australia
{gpopek,rkowalczyk}@groupwise.swin.edu.au

² Institute of Informatics, Wrocław University of Technology, Wrocław, Poland
{grzegorz.popek,radoslaw.katarzyniak}@pwr.wroc.pl

Abstract. Multiple approaches aim at describing numerical data with words. In this paper we give a brief overview of a distributed agent-based system providing summaries of city traffic in a textual form. The system deals locally with a problem of incomplete data adopting a method for grounding of modal statements in artificial cognitive agents. The internal uncertainty of an agent about a current state of the traffic is expressed with autoepistemic modal operators of possibility, belief, and knowledge.

Keywords: intelligent agent, incomplete knowledge, textual summaries of data.

1 Introduction

Distributed sensor networks are widely applied to gather data for various applications including habitat monitoring, wildfire monitoring, indoor localisation, navigation, and city traffic monitoring. A distributed nature of such networks, amounts of gathered data, and a general purpose of a system built on top of the sensor net all play a significant role when designing the flow and storage of data.

The main example assumed in this paper is a system providing linguistic descriptions of city traffic to drivers. In such a system the most precise data is needed only within local nodes of the network. Summaries of bigger areas can be performed based on aggregated data from single nodes.

There are existing approaches for assigning data with linguistic descriptions, especially in an area of Natural Language Generation (NLG). One of the most referred examples is an automatic generation of weather forecasts based on NLG system SUMTIME-MOUSAM [10], where main factors of a successful summary are its precision, short length, and lack of ambiguity. Authors of that system pay a lot of attention to a choice of words and the final form of the provided summary. Different approaches found in an area of Zadehian linguistic variables [13] introduce structured ways of processing summaries and aggregating them [4,5].

The key issue considered in this paper is a way in which the proposed system locally handles incomplete data and how it aggregates the data with respect

^{*} This paper was partially supported by Grant no. N N519 444939 funded by Polish Ministry of Science and Higher Education (2010-2013).

to provided patterns. On a local level in a traffic scenario each crossroad is monitored by an agent equipped with a set of sensors. The agent collects the data and generates local summaries. In cases of incomplete observation the agents generates summaries with auto-epistemic modal operators of possibility, belief, and knowledge by extending a solution described in [6,7]. Local summaries are further used to generate summaries for the region.

In Section 2 we discuss an application scenario and a general structure of the proposed system. In Section 3 we show how local summaries can be constructed; we provide an algorithm and present a scenario-based example. In Section 4 we provide a method for an aggregation of local summaries.

2 Current Approaches and Application Scenario

Traffic summaries are already provided in some form, e.g. in major cities of Poland newspaper sellers located at the main crossroads report recent changes in traffic density. Such an approach can bring many unwanted factors, e.g. observers change over time, they are not traffic experts, they are subjective, they lack context specific knowledge (e.g. about light cycle rotations).

There is an existing progress in designing and implementing automated solutions (e.g., VicRoads¹ measures car volumes passing through key road link in Melbourne and provides brief traffic descriptions for supported links; Google tracks ‘enabled’ devices and shows graphical summaries of traffic flow for areas where enough data is available) which reduce many of aforementioned problems related to a manual approach. Especially, an application of smart sensors can provide information about many factors of a traffic state (e.g. a number of vehicles located at each lane [1], the speed of passing vehicles and directions of a traffic flow [8], a classification of vehicles’ types [11]) in an automated way.

2.1 Application Scenario

Data about the numbers of cars flowing through main nodes of the road network is provided systematically after each 15-minutes interval. The goal is to provide an external user with a textual summary of the current traffic state in a queried ‘area’, e.g., a single junction, a district, or a more complex structure like the precise route. The system receiving data from sensors needs to store and process it in order to provide such a functionality to the user.

The system is supposed to be distributed. Each node should be able to locally process gathered data and provide a summary for the respective area using contextual knowledge (e.g. the length of a link, the size of a crossroad, their maximum throughput). Depending on the final architecture, nodes should be also capable of aggregating local summaries from a given set of surrounding (or underlying) nodes.

¹ VicRoads assists the Australian Government to achieve its integrated transport policy objectives in State of Victoria, AU.

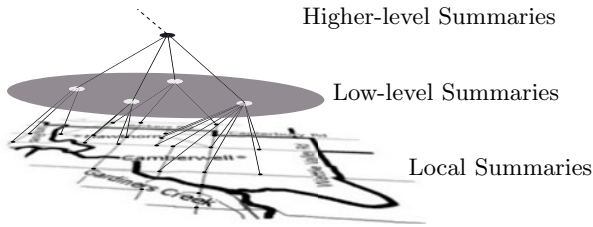


Fig. 1. Processing of summaries within a hierarchical architecture

In the most basic hierarchical aggregation (see Figure [1](#)) a user-query submitted via textual message can be parsed to the respective node which performs processing and a textual summary is sent back to the user. Another, more decentralised possibility is to make particular nodes more independent and regularly broadcast current summaries to surrounding nodes. Each node afterwards aggregates an up-to-date summary of its neighbourhood and supplies it to surrounding cars. Regardless of a structure of the aggregation the basic functionality of each node remains the same. It needs to aggregate summaries obtained from other nodes to obtain a summary.

The system needs to deal with partially missing data (due to sensors' failures or other types of measurement errors). The idea is not to use an imputation procedure (e.g. approximate or interpolate the missing data), but to reflect an incompleteness of data in the final summary. With this, a fact that a piece of information comes from a reasoning procedure rather than from a direct measurement of a traffic state can be preserved, providing more representative description of the situation with an indication of quality/completeness of observation data.

The main idea is to implement nodes as cognitive agents with history as their sets of experiences and — by extending a theory of grounding of modal statements of communication [\[6\]](#) to work within the provided model — to construct relevant local summaries using epistemic operators of Possibility, Belief and Knowledge. Summaries for more complex areas are to be obtained through aggregation of multiple local summaries.

3 Local Summaries

A basic preprocessing in terms of assigning linguistic-like labels to the traffic-state data is already performed (e.g., VicRoads in their visualisation assigns green, yellow or red colour to a link which stands for light, medium, and heavy traffic, respectively) based on a local context (e.g. an average volume of cars for a given link).

For the simplicity of the presentation it can be assumed, that data locally stored by an agent is expressed in that way. Therefore, each underlying object (a road link, a crossroad, a lane) is described with one property *Traffic* taking values from the set $\{light, medium, heavy\}$ or *null* when the value is missing.

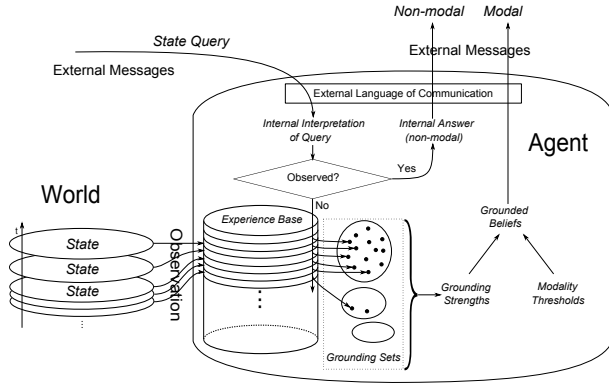


Fig. 2. A general structure of the local solution

Each object is assigned with its linguistic description (e.g. *at Camberwell Junction moving east into Riversdale Road*) which is used to fill the traffic description pattern “(belief) there is a (traffic) (object_text)” to obtain a textual summary (e.g., *I believe there is medium traffic going north into Glenferrie Road at the crossroad with Burwood Road*). In a situation when a respective piece of data is missing from the current observation the belief represented with an autoepistemic modal operator needs to be evaluated.

Johnson-Laird addresses [3] a possibility of use of a model theory to deal with modal reasoning. Dennett [2] states that “*exposure to x – that is, sensory confrontation with x over suitable period of time – is the normally sufficient condition for knowing (or having true beliefs) about x*”. Katarzyniak [7] uses mental models to support a choice of a modal statement related to a state of a chosen property in an object of an external environment in case where an actual observation regarding the state of the property is missing.

In the assumed approach [7] the agent in presence of missing data reduces its lack of knowledge using its previous experiences. One of natural assumptions accepted in a theory of grounding is an ability of natural cognitive agents to fill gaps experienced in an autonomously created model for actual worlds with some mental ‘patterns’ extracted from previous empirical experiences. Pieces of experience connected to concurrent states of the ‘gap’ are gathered in disjunctive grounding sets and lead to a raise of multiple alternative models. Each of these models can be assigned with a grounding strength with which it influences subjective convictions of the agent. Based on a distribution of these strengths and based on a system of modality thresholds (presented further) an agent grounds a certain set of statements using epistemic modal operators of possibility, belief, and knowledge.

As shown in Fig. 2 the agent receives a query. In case of an incomplete observational data the agent describes its cognitive attitude toward a state of unobserved property in a given object from an environment. Epistemic operators of possibility, belief and knowledge are used to reflect an agent’s internal state. Semantic

messages based on such epistemic operators (constructed according to rational restrictions) are assumed to be naturally understood by an external user of the system.

3.1 Grounding of Modal Statements

In a described scenario a property *Traffic* can exhibit one out of three possible values $\{light, medium, heavy\}$. The original model is based only on two values and an extension needs to be performed. In general it can be stated that for each object \circ and each property φ there are multiple possible states $\mathbf{u}_{\varphi,r}$, $r = \overline{1, r(\varphi)}$ for which corresponding modal statements $Pos(\varphi = \mathbf{u}_{\varphi,r})(\circ)$, $Bel(\varphi = \mathbf{u}_{\varphi,r})(\circ)$, $Know(\varphi = \mathbf{u}_{\varphi,r})(\circ)$ can be formed. Forming and choice of a correct formula should be done according to a relation of epistemic satisfaction for a given case and a system of modality thresholds constrained with a detailed set of requirements.

As for each property φ there are $r(\varphi)$ complementary states of a given object \circ and therefore, there is a set of $r(\varphi)$ different internal models (called mental models) representing a binding between formulas $(\varphi = \mathbf{u}_{\varphi,r})(\circ)$ and their meaning. In a proposed approach, each mental model is related to a certain state $\mathbf{u}_{\varphi,r}$ of a property φ and is induced by all those pieces of agent's experience, in which an agent observed an object \circ as exhibiting a state $\mathbf{u}_{\varphi,r}$ of a property φ .

The strength of a mental model is built through a positive correlation observed by an agent in its past. Each piece of experience, in which an agent observed an object \circ as exhibiting a state $\mathbf{u}_{\varphi,r}$ of a property φ , makes a corresponding mental model stronger in relation to the complementary models. We call this strength a grounding strength and for a given object \circ , a property φ , a moment t and for each $r = \overline{1, r(\varphi)}$ we denote it as $G_r(t, \circ, \varphi)$.

In an original approach [7] each grounding strength is evaluated as a cardinality of a set of supporting experience (i.e., number of all past situations in which an object exhibited a corresponding value of the property). This evaluation can be extended to include context dependency and, for example, to assign higher influence to those pieces of experience which are more similar to the current state of the environment (see [9]). Strategies for weighting and/or cutting irrelevant pieces of experience (including but not limited to: cardinality-based methods, semantic similarity/distance-based methods, reduction of grounding sets) can be directly applied in the proposed system.

3.2 Relation of Epistemic Satisfaction

The so-called relation of epistemic satisfaction describes conditions which have to be fulfilled by an agent's knowledge state (summarized in a form of grounding strengths) in order to make a modal statement grounded. A proposed theory defines a system of modality thresholds $\langle \lambda_{\min Pos}, \lambda_{\max Pos}, \lambda_{\min Bel}, \lambda_{\max Bel} \rangle$ and discusses in detail dependencies between $\lambda_{\min Pos}$, $\lambda_{\max Pos}$, $\lambda_{\min Bel}$, and $\lambda_{\max Bel}$ which need to be fulfilled in order to guarantee a rational language behaviour of an

agent. We propose a following definition of a relation of epistemic satisfaction adapted to our scenario:

Definition 1. *Relation of Epistemic Satisfaction for Modal Statements.* Let $\lambda_r(t, \mathfrak{o}, \varphi) = \frac{G_r(t, \mathfrak{o}, \varphi)}{\sum_{s=1}^{r(\varphi)} G_s(t, \mathfrak{o}, \varphi)}$. For a given system of modality thresholds, a moment t , an object \mathfrak{o} , a property φ and for each of its values $\mathbf{u}_{\varphi, r}$

- *Know*($\varphi = \mathbf{u}_{\varphi, r}$)(\mathfrak{o}) is grounded iff. $\lambda_r(t, \mathfrak{o}, \varphi) = 1$,
- *Bel*($\varphi = \mathbf{u}_{\varphi, r}$)(\mathfrak{o}) is grounded iff. $\lambda_{\min\text{Bel}} \leq \lambda_r(t, \mathfrak{o}, \varphi) < \lambda_{\max\text{Bel}}$,
- *Pos*($\varphi = \mathbf{u}_{\varphi, r}$)(\mathfrak{o}) is grounded iff. $\lambda_{\min\text{Pos}} \leq \lambda_r(t, \mathfrak{o}, \varphi) < \lambda_{\max\text{Pos}}$.

To preserve a rational behaviour of the system (not discussed here in detail) a system of modality thresholds needs to fulfil following inequalities:

$$\left\{ \begin{array}{l} 0 < \lambda_{\min\text{Pos}} < \lambda_{\max\text{Pos}} \leq \lambda_{\min\text{Bel}} < \lambda_{\max\text{Bel}} \leq 1 \\ \lambda_{\max\text{Pos}}, \lambda_{\min\text{Bel}} > 0.5 \\ \lambda_{\min\text{Pos}} \leq \min \left\{ \frac{1}{\eta}, \frac{1 - \lambda_{\min\text{Bel}}}{\eta - 1} \right\} \end{array} \right.$$

A parameter η can be related to limitations of focus of human attention and conscious processing. In context of above inequalities it guarantees the possibility of grounding at least η different *Pos* formulas at the same time.

Processing of a query is performed according to a structure of the approach presented in Figure 2. The whole process is initiated when an agent receives a query of a form $(\varphi = ?)(\mathfrak{o})$. It grounds possible answers according to an algorithm presented in Figure 3. The main advantage of providing message in such a form is an information given to the user that obtained summary does not directly come from an empirical measurement of described object, but is a result of a reasoning procedure.

3.3 Example

In the proposed system of textual summaries implemented with MATLAB (R2010a) each link is described with *id*, its capacity and textual representation for the summary (e.g., *at Camberwell Junction moving east into Riversdale Road*). Assume a crossroad of Glenferrie Road and Riversdale Road with links *r4402*, *r14642*, *r14520*, *r4274*. An example of generation of a local summary involves the situation of structure presented in Figure 4.

In this case a set of objects $\{r4402, 414642, r14520, r4274\}$ is described with properties $\{Interval, Day, Volume\}$. Since first two properties do not vary between objects (and are not a subject of summaries), they are presented (and implemented) in a compact way. An excerpt of compacted historical traffic data from October of 2006 is shown in Table 1.

If an agent has data, a simple pattern is filled to prepare a final summary in a linguistic form “(belief) there is a (traffic) (object_text)”. E.g., a summary

Require: State query $(\varphi = ?)(\circ)$

Ensure: Set of grounded formulas

```

1: if property  $\varphi$  in object  $\circ$  is observed and is equal to  $u_{\varphi,r}$  then
2:   Add  $(\varphi = u_{\varphi,r})(\circ)$  to grounded.
3: else
4:   for all  $r \in \{1, 2, \dots, r(\varphi)\}$  do
5:     Calculate  $\lambda_r(t, \circ, \varphi)$ .
6:     if  $\lambda_r(t, \circ, \varphi) = 1$  then
7:       Add  $Know(\varphi = u_{\varphi,r})(\circ)$  to grounded.
8:     end if
9:     if  $\lambda_{\min\text{Bel}} \leq \lambda_r(t, \circ, \varphi) < \lambda_{\max\text{Bel}}$  then
10:      Add  $Bel(\varphi = u_{\varphi,r})(\circ)$  to grounded.
11:    end if
12:    if  $\lambda_{\min\text{Pos}} \leq \lambda_r(t, \circ, \varphi) < \lambda_{\max\text{Pos}}$  then
13:      Add  $Pos(\varphi = u_{\varphi,r})(\circ)$  to grounded.
14:    end if
15:  end for
16: end if

```

Fig. 3. Generation of grounded statements

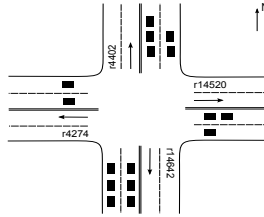


Fig. 4. A structure of a simple crossroad

based on interval t_n for link $r4402$ would be: *There is medium traffic going north into Glenferrie Road at the crossroad with Riversdale Road.*

To reply to a query about a volume for an object $r4402$ in the observation related to t_n the algorithm for generation of belief statements is invoked. The system of modality thresholds used in the traffic scenario has been set to $\langle 0.2, 0.6, 0.6, 1 \rangle$. Since description is related to Monday at 6:45-7:00, grounding experience is limited to weekdays and intervals from 6:45 to 7:15 (a way of limiting grounding experience needs to be provided by an expert for particular applications).

After evaluation $\lambda_{\text{Light}}(t_n, r4402, \text{Traffic}) = 0.15$, $\lambda_{\text{Medium}}(t_n, r4402, \text{Traffic}) = 0.85$ and $\lambda_{\text{Heavy}}(t_n, r4402, \text{Traffic}) = 0$ resulting in a set of beliefs: $Pos(\text{Traffic} = \text{Light})(r4402)$, $Bel(\text{Traffic} = \text{Medium})(r4402)$. One of the corresponding external messages in this situation takes a form: *I believe that there is medium traffic going north into Glenferrie Road at the crossroad with Riversdale Road.*

Table 1. Structure of data (October, 30th, 2006)

Time Index	Interval	Day	Volume			
			r4402	r14642	r14520	r4274
t_n	6:45-7:00	Monday	light	?	medium	light
t_{n-1}	6:30-6:45	Monday	light	medium	medium	light
t_{n-2}	6:15-6:30	Monday	light	light	medium	light
.

4 Aggregation

An aggregation of summaries about values of a property φ is performed according to predefined patterns. Patterns are simply sets of objects with an assigned textual description of the whole pattern. Within a traffic scenario it allows to model following types of traffic aggregations: areas at different scale, e.g. Boroondara, East Melbourne; precise routes; complex patterns, e.g. traffic leaving MCG.

An aggregated description for a given pattern $\mathfrak{P} = \{\circ_1, \circ_2, \dots, \circ_N\}$ is evaluated based on weights $w(\circ_n)$ of objects within pattern (e.g., in our scenario weights are highly correlated with an average flow through the link), local summaries of objects and a system of modality thresholds.

An initial aggregation of a property φ for a pattern \mathfrak{P} is given by

$$\langle w(\mathfrak{P}, \mathbf{u}_{\varphi,1}), w(\mathfrak{P}, \mathbf{u}_{\varphi,2}), \dots, w(\mathfrak{P}, \mathbf{u}_{\varphi,r(\varphi)}) \rangle,$$

where each $w(\mathfrak{P}, \mathbf{u}_{\varphi,r})$ is a sum of modified weights of those objects which were described in local summaries as exhibiting a value $\mathbf{u}_{\varphi,r}$ of a property φ . The modified weight is equal to:

- $w(\circ)$ if no modality or a modality *Know* is present in the local summary,
- $\lambda_{\min\text{Bel}} \cdot w(\circ)$ if a modality *Bel* is present in the local summary,
- $\lambda_{\min\text{Pos}} \cdot w(\circ)$ if a modality *Pos* is present in the local summary.

Modified weights reflect a lowered impact on the final outcome of those local summaries which come from an incomplete knowledge. In above formulas lowest modality thresholds for respective modal operators are used to modify their impact. Used values depend on a particular application however they should be chosen from intervals related to respective modalities.

There are two main steps in the process of an aggregation algorithm depicted on Figure 5. After an evaluation of an initial aggregation in steps 1-3, the main choice of an index r^* of an output value is done in step 4. An initial aggregation (after normalization) is a subject to a defuzzification method [12] which needs to be provided by an expert depending on an application area. In the provided example an element with a highest aggregated weight is chosen.

A parameter w^* evaluated in step 6 reflects a degree of uncertainty received from local nodes. It gets lower with a rising number of missing observations in local nodes which result in higher number of local summaries being modal.

Require: State query $(\varphi = ?)(\mathfrak{P})$, local summaries, system of modality thresholds.

Ensure: A final summary.

```

1: for all  $r \in \{1, 2, \dots, r(\varphi)\}$  do
2:   Calculate  $w(\mathfrak{P}, \mathbf{u}_{\varphi, r})$ .
3: end for
4: Assign  $r^* := \text{argmax}(w(\mathfrak{P}, \mathbf{u}_{\varphi, r}))$ 
5: Assign  $w^* := \left( \sum_{r \in \{1, 2, \dots, r(\varphi)\}} w(\mathfrak{P}, \mathbf{u}_{\varphi, r}) \right) / \left( \sum_{\circ \in \mathfrak{P}} w(\circ) \right)$ 
6: if  $w^* = 1$  then
7:   Add  $(\varphi = \mathbf{u}_{\varphi, r})(\mathfrak{P})$  to a summary.
8: end if
9: if  $\lambda_{\min\text{Bel}} \leq w^* < \lambda_{\max\text{Bel}}$  then
10:  Add  $\text{Bel}(\varphi = \mathbf{u}_{\varphi, r})(\mathfrak{P})$  to a summary.
11: end if
12: if  $\lambda_{\min\text{Pos}} \leq w^* < \lambda_{\max\text{Pos}}$  then
13:  Add  $\text{Pos}(\varphi = \mathbf{u}_{\varphi, r})(\mathfrak{P})$  to a summary.
14: end if

```

Fig. 5. Generation of grounded statements

Values closer to 1 reflect a higher impact of directly gathered data about the current state of the network. Final steps 6-14 are generating a summary based on w^* and a system of modality thresholds.

Generation of an output textual summary is analogical as in a case of generation of a textual local summary as each pattern is assigned with its textual description. An example of an aggregated summary would be: *I believe there is a heavy traffic in Boroondara.*

5 Summary

We proposed a distributed multi-agent system capable of a generation of textual summaries of a partially observed environment. The system provides summaries involving auto-epistemic operators of possibility, belief, and knowledge in situations where data is incomplete. In such cases the internal reasoning procedure is performed in order to evaluate the summary. Summaries are generated in a natural language form in order to make them comprehensible by a non-expert external user of the system. This enhances the current approaches that either use simple indicators (e.g. colours) based on the local counts and/or do not reflect the fact of the quality of observation data (e.g. missing data) in their summaries.

An algorithm for the evaluation of local summaries based on the theory of grounding is proposed. The method for an aggregation of local summaries according to predefined patterns provided by a domain expert is described and supported with an algorithm. The system in its current form is designed to work with data provided in form of objects exhibiting discrete-valued properties. In the provided scenario we have used a predefined way of translation of measured

car volumes into linguistic labels *low*, *medium*, *heavy*. Each one of these labels is activated exclusively at a given point of time.

The next research step is extend the model and treat values of properties present in the system as zadehian fuzzy-linguistic labels [13] with degrees of membership and perform a processing by applying and adapting solutions from the fuzzy-linguistic domain.

References

1. Bas, E., Tekalp, A.M., Salman, F.S.: Automatic Vehicle Counting from Video for Traffic Flow Analysis. In: 2007 IEEE Intelligent Vehicles Symposium, June 2007, pp. 392–397 (2007)
2. Dennett, D.C.: True believers: The intentional strategy and why it works. In: Stich, Warfield (eds.) *Mental Representation: A Reader*. Blackwell (1994)
3. Bell, V.A., Johnson-Laird, P.N.: A model theory of modal reasoning. *Cognitive Science* 22(1), 25–51 (1998)
4. Glöckner, I.: Optimal selection of proportional bounding quantifiers in linguistic data summarization. In: *Soft Methods for Integrated Uncertainty Modelling. Advances in Soft Computing*, vol. 37, pp. 173–181. Springer, Heidelberg
5. Kacprzyk, J., Wilbik, A.: Linguistic Summarization of Time Series Using Fuzzy Logic with Linguistic Quantifiers: A Truth and Specificity Based Approach. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) *ICAISC 2008. LNCS (LNAI)*, vol. 5097, pp. 241–252. Springer, Heidelberg (2008)
6. Katarzyniak, R.: The language grounding problem and its relation to the internal structure of cognitive agents. *Journal of Universal Computer Science* 11(2), 357–374 (2005)
7. Katarzyniak, R.: On some properties of grounding uniform sets of modal conjunctions. *Journal of Intelligent and Fuzzy Systems* 17(3), 209–218 (2006)
8. Lee, J., Bovik, A.C.: Estimation and analysis of urban traffic flow. In: *IEEE International Conference on Image Processing (ICIP 2009)*, pp. 1157–1160 (2009)
9. Popek, G.: Strength of Formula Grounding and Context-Dependent Strategies for Grounding Experience Determination. In: *Challenging Problems of Science, Comp. Sci., Knowledge Processing and Reasoning for Inf. Society*, pp. 21–38. APH EXIT (2008)
10. Reiter, E., Sripada, S., Hunter, J., Yu, J., Davy, I.: Choosing Words in Computer-Generated Weather Forecasts. *Artificial Intelligence* 67, 137–169 (2005)
11. Tuan, H.T., Kostia, R., Sijun, L., Jian, Z.: Vehicle Classification at Nighttime Using Eigenspaces and Support Vector Machine. In: *2008 Congress on Image and Signal Processing*, May 2008, pp. 422–426 (2008)
12. Yager, R., Filev, D.: *Essentials of Fuzzy Modeling and Control*. Wiley, NY (1994)
13. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning-I. *Inf. Sci.* 8(3), 199–249 (1975)

Target Factor Algorithm of PSC Ship-Selecting System Based on Rough Set Theory and Hierarchic Analysis

Zhonghua Sun and Tingting Yang

College of Navigation, Dalian Maritime University, Dalian 116026, China
yangtingting820523@163.com, zhonghua9669@163.com

Abstract. This paper research Paris-MOU NIR (New Inspection Regime) of latest PSC rules of 2009/16/EC, Combining the rough set theory with hierarchic analysis model in view of not reduce efficient classify information, introducing the definition of importance into non-core attributes, a new data reduction algorithm is proposed and put forward an algorithm of PSC ship-selecting system. The results reveal that the algorithm is able to resolve data reduction problem and simplify network structure.

Keywords: rough set theory, hierarchic analysis, data reduction.

1 Introduction

2009/16/EC are the newest European PSC(Port State Control) rules published by European Congress and committee in May, 2009. It emphasis on an agreement on ship risk assessment system between Paris-MOU, to build up a scientific check procedure and retention management called NIR (New Inspection Regime) [1].

In order to assess ship's safety risk reasonably, NIR set up risk ranks calculated by risk models. The model sufficiently consider recent research performance of IMO, and divide risk factor to two types. One is based on general parameters, such as ship type, ship age, ship flag state, recognizing organization and ISM company management performance. The other is history parameters, such as retention times and defect numbers in a period. But the ship-selecting algorithm just select some representative factors, and simply add each risk value which is scored according to risk factor. They are not comprehensive and scientific, because ship-selecting model should be a non-linear but not linear one. Wei[2] analyzes the main risk factors of ships and introduces the BP neural network in setting up the model of ship selection for FSC inspection. Zhou[3] introduces and explains the mathematic background of 'Black, Gray and White' Flag State grading system which has been widely used in some regional organizations, and suggests using the same mathematic principle to solve the periodical Matthew Effect and discontinuity problem, by introducing Excess Factor(EF) as revision of the targeting system.. But they not solve the problem of selecting scientific target factor.

In this paper, 2009/16/EC NIR is set as research object, combining the rough set theory with hierarchic analysis model in view of not reduce efficient classify information, introducing the definition of importance into non-core attributes, a new data reduction algorithm is proposed and put forward an algorithm of PSC ship-selecting

system. The algorithm is able to resolve data reduction problem and simplify network structure.

2 Data Reduction Based on Rough Set Theory and Hierarchic Analysis

Rough set theory is firstly proposed by a Poland mathematician Pawlak in 1982. It is a theory used to deal with uncertainly, non-concise and incomplete data. Rough set theory method can deduce decision or classify regulation through knowledge reduction in the premise of keeping classify capability. Because of knowledge deduction based on rough set theory not need give some characteristics and related knowledge in advance, we can get knowledge reduction and simplification from knowledge now available. It's a new mathematical method for dealing with uncertainly, non-concise, incomplete and inconformity information and knowledge, so rough set theory become a powerful knowledge obtaining tool for machine learning, model identification, decision analysis, control process, data mining and specialist system, especially domain of knowledge mining for data base.

Hierarchic analysis method is proposed by an American Operational Research professor T. L. Satty in the middle 1970s. The method and procedure for decision problem is: firstly analysis relationship between inherent factor through sufficiently research, then divide it for some layers including Indicator layer, Criteria layer (containing subsidiary criteria layer) and Goal layer, etc.

For one knowledge base, corresponding decision table is likely to have multi-attribute reduction. Due to the purpose of knowledge reduction is produce decision rules for decision table, then the attribute numbers can directly affect complexity degree of decision rules, so we all want to get the least number of it. In order to avoid generate the same minimum attribute reduction when just make use of rough set theory algorithm, and also avoid appear more subjective composition, here we combine hierarchic analysis method with rough set theory to advance rationality of reduction. The specific attribute reduction algorithm is:

(1) Build Up Simple Dissimilar Matrix

Build up simple dissimilar matrix can judge uniformity of decision table, and divide non-uniformity decision table into two sub-tables, including uniformity decision table and non-uniformity decision table which is classified according to attributes. The specific algorithm is: firstly, find out pre-deleted items that cause class non-uniform; then delete relative non-primary classify parameter; the last parameter can constitute reduced evaluation parameter set. The algorithm process is described as[4]:

We hypothesize the sample set $S = \{X_1, X_2, \dots, X_n\}$, n is total number of the sample; evaluation parameter set $C = \{C_1, C_2, \dots, C_m\}$, m is evaluation item number; ship risk rank

$D = \{d_1, d_2, \dots, d_k\}$, k is classify number. The evaluation value of sample X_i is $VC_j(X_i)$, it's ship risk rank is $VD(X_i)$.

Put in: a sample data set S , each sample contain m evaluate parameter and a risk class d

Put out: reduced evaluation parameter set Y.

Step1: Refer to 2009/16/EC NIR ship rank evaluation system, convert each sample evaluation parameter u into relative rank category C.

for i=1 to n do

for j=1 to m do {uj (Xi) -- >Cj (Xi) }

Step2. Find out pre-reduced parameter, then put it into dtable.

for j=1 to n do {i=VD (Xj) ; Xj→table (i) ; } % classify sample according to category

for i=1 to k do %find out evaluation parameter which has the same category but the value of parameter item is different

for j=1 to m do

{All the sample of table (i) : if the category of Cj item is different, put (Cj, di) →dtable which is to be reduced ; }

Step3 Parameter reduction

While dtable is not \emptyset , do

{take out a log from table, parameter to be reduced noted as Cx, category noted as dx ;

bj (Cx) = true ; % bj is noted as has been reduced or not

for j=1 to n do % if Cx is not primary classify parameter, it can be deleted

{if VD (Xj) < > dx

if VCx (Xj) > = VD (Xj) {bj (Cx) = false ; break ; } }

if bj (Cx) = false clear away the log from dtable ; get out the next log from dtable ; }

Step4. If table is not \emptyset , get away evaluation parameter according to each log, then it will be the evaluation parameter set after reduction.

(2) Calculate Core Set H of Decision Table

(3) Put Non-core Attribute into Core Set

The next step is get importance of non-core attribute relative to decision attribute, then get the reduction after putting non-core attribute into core set according to importance high to low. When we want to get importance of non-core attribute, put it as Indicator layer, and put the comparative largest importance of decision attribute as Goal layer. Here, we need refer to the opinion of specialist to determine Criteria layer and the relationship of comparative importance.

The specific algorithm of adding non-core attribute is:

Confirm relative importance of Criteria layer and non-core attribute.

Make use of hierarchic analysis model to get relative importance of non-core attribute to decision attribute. And hypothesize it as {c1,c2,...,cm} according to importance high to low.

Put {c1},{c1,c2},...,{c1,c2,...,cm} into it in order, check out the consistency of decision table. If is consistent, when add to {c1,c2,...,ck}, we get the reduction of decision table, it's {c1,c2,...,ck}+H.

(4) Get the Optimum Reduction

Delete one non-core attribute of the reduction from step (3) at random, rebuild decision table and simple dissimilar matrix, and check out non-core attribute in order. If after reduced , it not change the consistency of decision table, we should reduce this attribute. Repeat step (4), the last attributes is the reduction.

3 The Application of Data Reduction Based on Rough Set Theory and Hierarchic Analysis in PSC Ship-Selecting System

In this paper, we comprehensively consider national and international primary ship-selecting regime of MOU, such as Paris-MOU, Tokyo-MOU, referring their ship-selecting factor to build up one decision table[5-6]. According to PSC check result of Liaoning Maritime Bureau during the period of 2006 to 2010, we set up a ship risk factor decision table. The Goal layer $G = \{\text{Detain}\}$ is decision attribute ; Criteria layer $B = \{\text{ship attribute, administration institute, history check}\}$; Indicator layer $C = \{\text{ship type, ship age, mariner, route, tonnage ; Flag state, recognized organization, company ISM system, ship ISM system, weather geography and hydrology ; defect number during the last 36 months check log., detained number during the last 36 months check log, duration from last check, duration from last detain, exist non-correct defect}\}$. Indicator A is detain rate, different route, whether build up ISM. We build up the following Hierarchic structure chart, see figure1. And Table 1 indicates the judging matrix of Criteria layer. Due to the limited length of article, here we just give an example of Criteria Layer. The specific calculate procedures are:

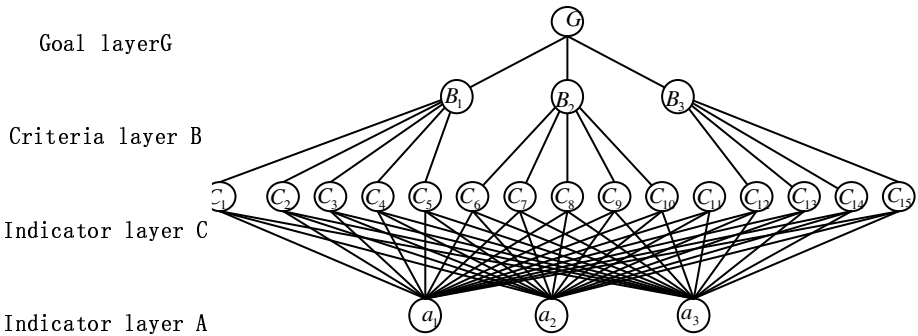


Fig. 1. Hierarchic structure chart

(1) Calculate element multipliers of judging matrix rows[7]

$$M_1 = 1 \times 1 \times 1 / 8 = 1 / 8, M_2 = 1 \times 1 \times 1 / 8 = 1 / 8, M_3 = 8 \times 8 \times 1 = 64$$

(2) Calculate n times square root of M_i

$$\bar{W}_1 = \sqrt[3]{M_1} = \sqrt[3]{1/8} = 0.5, \bar{W}_2 = \sqrt[3]{M_2} = \sqrt[3]{1/8} = 0.5, \bar{W}_3 = \sqrt[3]{M_3} = \sqrt[3]{64} = 4$$

Table 1. Judging matrix of Criteria layer

	Ship attribute B_1	Administration B_2	institute B_3
Ship attribute B_1	1	1	1/8
Administration institute B_2	1	1	1/8
History check B_3	8	8	1

$$(3) \text{ Regularization vector } \bar{W} = [\bar{W}_1, \bar{W}_2, \bar{W}_3]^T = [0.5, 0.5, 4]^T$$

$$\sum_{j=1}^n \bar{W}_j = 0.5 + 0.5 + 4 = 5, \quad W_1 = \frac{\bar{W}_1}{\sum_{j=1}^n \bar{W}_j} = 0.5/5 = 0.1$$

$$W_2 = \frac{\bar{W}_2}{\sum_{j=1}^n \bar{W}_j} = 0.5/5 = 0.1, \quad W_3 = \frac{\bar{W}_3}{\sum_{j=1}^n \bar{W}_j} = 4/5 = 0.8$$

Then, the wanted Characteristic vector $W = [0.1, 0.1, 0.8]^T$

(4) Calculate the maximum Characteristic Values of judging matrix

$$AW = \begin{bmatrix} 1 & 1 & 1/8 \\ 1 & 1 & 1/8 \\ 8 & 8 & 1 \end{bmatrix} \times \begin{bmatrix} 0.1 \\ 0.1 \\ 0.8 \end{bmatrix}$$

$$(AW)_1 = 1 \times 0.1 + 1 \times 0.1 + (1/8) \times 0.8 = 0.3, \quad (AW)_2 = 0.3, \quad (AW)_3 = 2.4$$

$$\lambda_{\max} = \sum_{i=1}^n \frac{(AW)_i}{nW_i} = \frac{(AW)_1}{3W_1} + \frac{(AW)_2}{3W_2} + \frac{(AW)_3}{3W_3} = 3$$

$$CI = \frac{\lambda_{\max} - n}{n-1} = \frac{3-3}{3-1} = 0, \quad CR = \frac{CI}{RI} = \frac{0}{0.52} = 0 < 0.1$$

This explains that it has good consistency.

In the same way, we also use hierarchic analysis method to get weight in order of Indicator layer A is (0.1524, 0.1375, 0.1142, 0.0833, 0.0126 ; 0.1207, 0.0872, 0.0424, 0.0384, 0.0113; 0.0582, 0.0516, 0.0463, 0.0202, 0.0237).

From the data above, we can conclude that the reduction of ship-selecting system target factor is {ship type, ship age, mariner, Flag state, recognized organization, route, defect number during the last 36 months check log, detained number during the last 36 months check log}.

Then we can use fuzzy comprehensive evaluation method to set up evaluation set and evaluation subordinate degree, confirm fuzzy calculator to build up mathematical evaluation model of ship-selecting system. If we just use rough set method to reduce data, it will emerge multi-reduction which may be affected by specialist's subjectivity.

4 Conclusion

This paper research Paris-MOU NIR (New Inspection Regime) of latest PSC rules of 2009/16/EC, Combining the rough set theory with hierarchic analysis method to propose a new algorithm of PSC ship-selecting system. The algorithm effectively solve the problem of attribute reduction, not only simplify network structure but also avoid the effect of specialist's subjectivity.

References

1. Yi, X.: NIR A new weapon of Paris-MOU. The Waterborne Safety. China Ship Survey (August 2009)
2. Wei, D., Chen, L.-L., Zeng, Q.-S., Qiu, H.-Z.: Research on modeling of ship selection for FSC inspection based on neural network. China Maritime Safety. Maritime Management (August 2010)
3. Zhou, C.: How to avoid Matthew Effect in selecting target vessels for Port State Control. In: China Maritime Safety. Maritime Workshop, December 2008, pp. 37–40.
4. Zhang, X.-F., Tian, X.-D., Zhang, Q.-L.: Data Reduction Based on Rough Set Theory and Hierarchic Analysis. Journal of Northeastern University(Natural Science) 29(1) (2010)
5. The Paris-MOU Secretariat. Annual Report Onport State Control In The Atlantic Region (EB/OL) (2010), <http://www.paris-mou.org>
6. Ningbo: Analysis of the new ship-targeting system of the Paris MOU. China Maritime. Expert's View, pp. 21–24 (September 2010)
7. Lou, H.: Research on PSC Targeting System And It's Implementation. Master thesis of Dalian Maritime University (October 2008)

Route Planning Based on Combination of Artificial Immune Algorithm and Ant Colony Algorithm

Qingfeng Wang and Yuhui Wang

School of Mechanical Engineering and Automation of Beihang University
Beijing 100191, China
wqfstudent@163.com, wangyuhui64611@buaa.edu.cn

Abstract. Artificial immune algorithm and ant colony algorithm are combined to deal with problem of 2D route planning of aircraft. Initial routes are generated randomly within the flying area and clonal selection algorithm is used to search good routes. A group of routes with minimum cost of threat and oil are gained. Some initial pheromone is put nearby these routes. Based on this, ant colony algorithm are used to search optimal route while threaten avoid and minimum cost are taken into consideration.

Keywords: route planning, artificial immune, ant colony algorithm, clonal selection, mutation.

1 Introduction

Route planning is to generate a minimum cost route of aircraft between start point and end point based on a predetermined cost function with certain constraints [1]. Several representative algorithms that have been applied for route planning are as follow.

A* algorithm is one of classical heuristic search algorithm. An improved A* algorithm was adapted to search the optimized route by S.Al-Hasan [2]. Evolutionary algorithm [3] has strong robustness to solve complex optimum problems. But precocious were found sometimes. Artificial neural networks algorithm [4] gets local minimization too easily and converges slowly. Ant colony algorithm is a new heuristic algorithm for route planning [5]. It is simple and easily-accomplished but is easy to fall in local optima. Genetic algorithm was used to obtain optimal routes of aircraft [6]. Particle swarm algorithm was used to search the best route of UAV [7].

Artificial immune algorithm(AIA) is a new computational intelligence approach. Combination of intelligent optimization algorithms can redeem a defect while optimization algorithm is used separately. Artificial immune algorithm can avoid "blindness selection" of ants and can make algorithm more efficient [8].

In this paper, AIA and ant colony algorithm are combined to deal with route planning of aircraft. First, AIA is employed to obtain several routes with low cost of oil and threat. Then ant colony algorithm is used to generate optimized route.

2 Problem of Route Planning

Route planning is to generate a minimum cost route of aircraft between start point and end point. In Fig 1., the inclined rectangular is area available for flying .The circles are represented threat areas that cannot fly in.

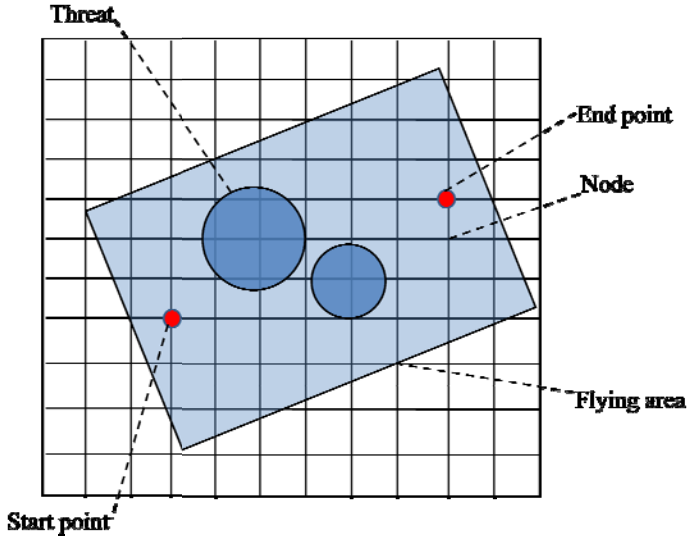


Fig. 1. Problem of flying area and route planning

Generalized cost function [9] is used to evaluate the route. It is given as follow.

$$W = \int_0^L [kw_t + (1 - k)w_f] dt \tag{1}$$

In formula (1), L is length of route, W is generalized cost_function, w_t is threat cost of route, w_f is oil cost of route and k is weight coefficient.

3 Generate the Best Group of Routes by AIA

3.1 Artificial Immune Algorithm and Related Concepts

AIA is one of the nature-inspired approaches in solving the optimization problem based on principle and mechanism of biological immune system. AIA have the advantage to prevent the population from being trapped into local optimum [10].

Some concepts of AIA for engineering is follow [11].

Antigen is the target or solution of the problem. Antibody is corresponding to candidate solution of the problem. Antibody population is a set of Antibodies. Affinity represents the matching degree between antibody and antigen.

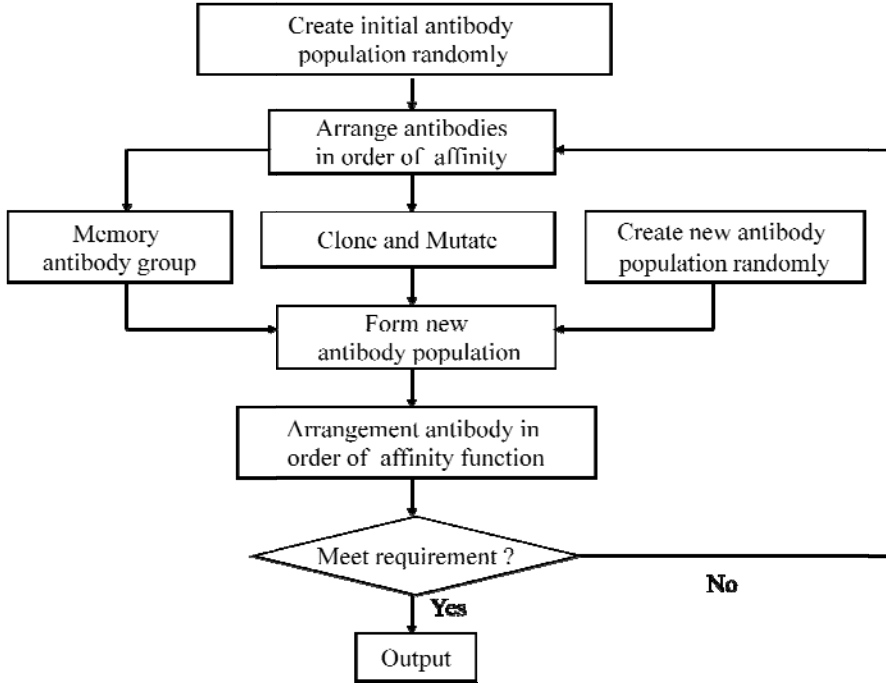


Fig. 2. Flow chart of obtaining the best group of routes by AIA

3.2 Generate the Best Group of Routes by AIA

In this paper, clonal selection algorithm [12], one of artificial immune algorithm, is adapted to search a group of preliminary routes with low cost. Fig 2 is the flow chart of obtaining the best group routes by clonal selection algorithm. The best group of routes will be obtained after iterations.

3.2.1 Generation of Initial Antibodies and Calculation of Affinities

Every antibody is a route candidate for the route planning within flying area. Antibody population consists of a certain amount of antibodies. Initial antibody population is created randomly as illustrated in Fig 3. Connect the start point and end point. Then a set of lines perpendicular to line SE is drawn. Each perpendicular line interests the planning area boundary line in two points $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$. A random data $\lambda_i (0 \leq \lambda_i \leq 1)$ is generated for every perpendicular line. The point $P_i (i = 1, 2, \dots, N)$ in this line is given as formula (2).

$$P_i = \lambda P_{i1} + (1 - \lambda)P_{i2} \tag{2}$$

Connect every point P_i in succession and add point S in begin E in the end. An antibody is constructed.

For every segment of an antibody, segmentation points with proportion of 0.1, 0.3, 0.5, 0.7 and 0.9 are gotten. The threat cost of an antibody is given by formula (3) [9].

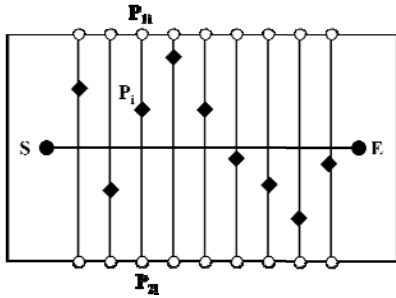


Fig. 3. Antibody point

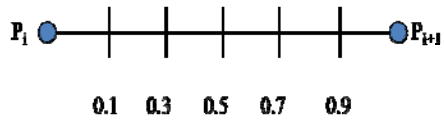


Fig. 4. Segmentation points

$$w_{t,i} = L_i \sum_{j=1}^N \left(\frac{1}{d_{0.1,i,j}^4} + \frac{1}{d_{0.3,i,j}^4} + \frac{1}{d_{0.5,i,j}^4} + \frac{1}{d_{0.7,i,j}^4} + \frac{1}{d_{0.9,i,j}^4} \right) \tag{3}$$

i is the number of segments in an antibody, j is number of threats, $d_{0.1,i,j}$ is the cost of threat of edge i caused by threats j at the point of $1/10$ length from start point. The meaning of $d_{0.3,i,j}$, $d_{0.5,i,j}$... $d_{0.9,i,j}$ is similar to that. Cost of oil is directly proportional to the length of route. The total cost is defined by formula (1) in section 1.

3.2.2 FCM Cluster Analysis to Get Best Cluster of Antibodies

First, arrange antibodies in order of cost. Then clustering analysis of antibodies according to affinities is implemented by FCM [13]. Because the antibody with maximum affinity corresponds to that of minimum cost, cost is taken as parameter of clustering analysis. Let k represent the sequence number of an antibody in antibody population. Hence W_k represent the cost of the k th antibody. The objective function of clustering is given by formula (4).

$$J_f = \sum_{j=1}^c \sum_{k=1}^n [\mu_j(W_k)]^b \|W_k - m_j\|^2 \tag{4}$$

$$\sum_{j=1}^c \mu_j(W_k) = 1 \quad , \quad k = 1, 2, \dots, n$$

In formula (4), m_j is the cost value of the j th cluster center, c is number of cluster, $\mu_j(W_k)$ is membership degree that W_k belongs to m_j fuzzily.

$$m_j = \frac{\sum_{k=1}^n [\mu_j(W_k)]^b W_k}{\sum_{k=1}^n [\mu_j(W_k)]^b}, j=1,2,\dots,c \quad (5)$$

$$\mu_j(W_k) = \frac{\left(1/\|W_k - m_j\|^2\right)^{1/(b-1)}}{\sum_{i=1}^c \left(1/\|W_k - m_i\|^2\right)^{1/(b-1)}}, k=1,2,\dots,n, j=1,2,\dots,c \quad (6)$$

The cluster of antibodies with minimum cost (maximum affinity) is gained after FCM cluster analysis. They are taken into memory antibody group and will be added into the next generation of antibody population. It is the first group for new antibody population.

3.2.3 Clone and Mutation

For the antibodies in the top 80 percents, clone is implemented. The problem of that route cannot through area of threat circles is not considered here. It will be consider in the step of ant colony algorithm. The number of clones is given by formula (7).

$$N_{ck} = \frac{\Phi \times N}{k + 1} \quad (7)$$

Φ is coefficient (varying with the amount of antibodies in the antibody population), N is the amount of antibodies and k is the sequence number of the antibody according to the cost.

For each antibody mutation rate is inversely proportional to its number of clone.

$$P_{mk} = \frac{10}{N_{ck} + 1} \quad (8)$$

In formula (8), p_{mk} is mutation rate of each antibody k , N_{ck} is number of clones of antibody with sequence number k .

Gaussian model [14] is adopted to deal with mutation of antibodies. Gaussian distribution is expressed as formula (9).

$$F_c(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty \quad (9)$$

In formula (9), μ is mean of a normal distribution, σ is standard deviation.

The implement method of antibodies mutation by Gaussian mutation is as follow.

$$\lambda'_i = \lambda_{ki} + N(0, \sigma) \quad (10)$$

$N(0,1)$ is the normal distribution. For each point P_i of the antibody with sequence number k , mutation rate is calculated by formula (10). A random data $v_i(0 \leq v_i \leq 1)$

is generated. If $v_i \leq p_{mk}$ P_i replace by a new point. The new point is recalculated by formula (2) with λ_i' calculated by formula (10).

All antibodies that take place mutation are put into a group as candidates of the next generation of antibody population. It is the second group for new antibody population.

The third group of antibodies is generated randomly in the way of generation of initial antibodies.

3.2.4 Gain the Best Group Routes

Three groups of antibody candidates are combined and are arranged according to cost. The best N antibodies are remained to form the new generation of antibody population while population diversity is taken into account. Iteration in the way display in Fig 2 is implemented repeatedly. Until the difference value of average cost of the best cluster between two times is within the given value, the best group of routes is obtained. They are correspondence of the antibodies in the best cluster.

3.2.5 Placement of Initial Pheromone

As show in Fig 1, the whole area is divided into certain amount of grids. Every point that a horizontal line and a vertical line cross in is taken as a node.

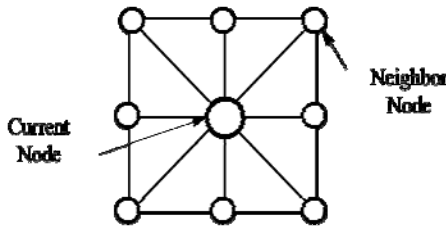


Fig. 5. Eight neighboring nodes of current node

Fig 5 shows the neighboring node of current node , each node has eight neighboring nodes. Each segment between two nodes is taken as an edge.

After best antibody group are gained. A group of edges nearest to the routes is found. Some initial pheromone is laid out in there edges for ants to find the best route quickly and to avoid the blindness in searching process.

4 Ant Colony Algorithm

For ant colony algorithm, first, lay out pheromone to every edge at the same amount. Each ant just can creep to one of its neighboring node. Which neighboring node will ant creep to due to the state transition probability as giving by formula (11). Each ant lay out certain amount of pheromone to the edge it just through. And after several times creep of ants the best route is gained.

$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_{ik}(t)]^\beta [r_{ij}(t)]^\gamma}{\sum_{s \in \text{Allowed}_k} [\tau_{is}(t)]^\alpha [\eta_{is}(t)]^\beta [r_{is}(t)]^\gamma} & \text{if } j \in \text{Allowed}_k \\ 0, & \text{other} \end{cases} \quad (11)$$

i is serial number of current node, j is serial number of candidate node, α is factor of pheromone, β is heuristic factor and $\eta_{ij}(t) = \frac{1}{d_{ij}}$ is heuristic function. The formula

[9] is improved by adding $[r_{ij}(t)]^\gamma$. It represents guidance to ants by end point. If a node is inside the threat circle, let $p_{ij}^k(t) = 0$. Then the ant cannot creep to it.

$$r_{ij}(t) = \chi d_j \quad (12)$$

d_j is the distance between candidate node and end point.

After every update of pheromone, the new amount of pheromone of each edge that connect two nodes with number of i and j is calculated by formula (13)

$$\tau_{ij}(t+n) = (1-\rho) \cdot \tau_{ij}(t) + \Delta\tau_{ij}(t) \quad (13)$$

In formula (13), ρ represents coefficient of pheromone volatilization, it is used to simulate volatilization of pheromone in nature.

5 Experiment and Discuss

5.1 Parameter and Experimental Result

Parameters listed in Table 1 are adapted for experiment. The average cost value of the best cluster at each iteration of clonal selection algorithm is illustrated in Fig 6. In Fig 7, fine polylines represent best group of routes after conal selection algorithm and the thick polyline is the best route after ant colony algorithm. The inclined rectangular is area available for route planning. The start point is in left while the end point is in the right. In Fig 8, contrast of the cost value of routes with clonal selection and that without clonal selection is displayed.

5.2 Discuss

The average cost of the best cluster decrease along with iterations times of clonal selection increasing and the best group of routes can be obtained at the end. Fig 8 indicates that the cost values of routes after each iteration of ant colony algorithm with clonal selection are small than that without clonal selection algorithm in one to one correspondence. This shows clonal selection algorithm is beneficial for ants to find the best route quickly.

Table 1. Parameters

parameter	Meaning of parameter	value
N	Antibody population size	80
c	Number of cluster	10
b	Cluster parameter	2
N_g	Ant Group Number...	10
N_a	Ant Number of a group.	10
N_p	Point Number of Antibody	11
k	Weight coefficient	0.9
α	Factor of pheromone	1.0
β	Heuristic factor g	1.2
γ	Guidance factor	80
ρ	Volatilization factor	0.1

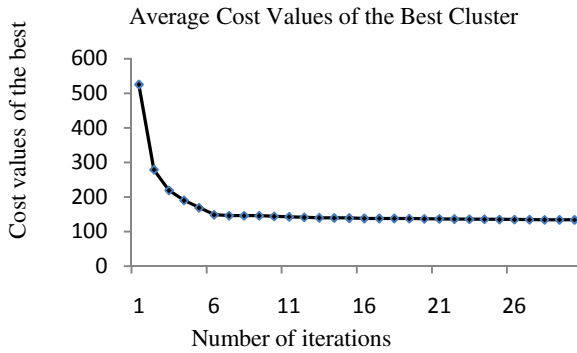


Fig. 6. The average cost of the best cluster

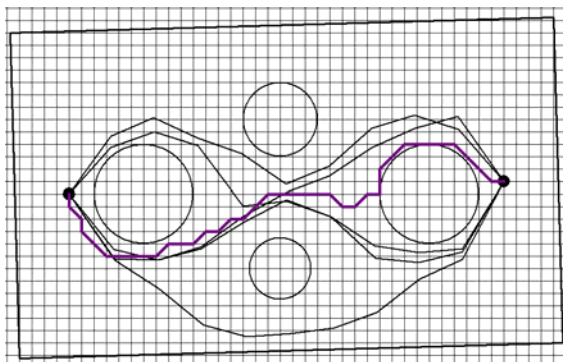


Fig. 7. The best group of routes obtained by AIA and the best route

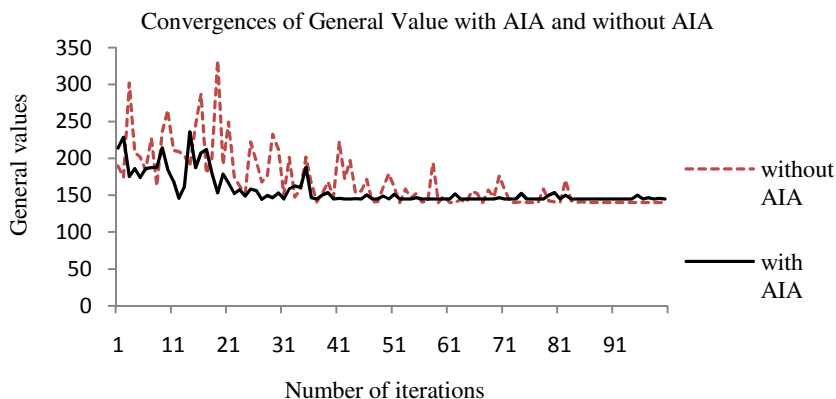


Fig. 8. Cost value of routes with AIA and without AIA

6 Conclusion

A method of combining AIA and ant colony algorithm is proposed. Clonal selection algorithm is used to search the best group of routes in initial phase can reduce the possibility of sinking into local optima. It can guides ants to find the best route quickly and to avoid the blindness. Combination of AIA and ant colony algorithm can improve the efficiency of route planning algorithm.

Acknowledgements. The work was supported by the Aviation Science Funds under Grant No. 20101751020.

References

1. Du, P., Yang, C.: Introduction of Air Vehicle Path Planning Algorithms. *Flight Dynamics* 23(2), 10–14 (2005)
2. Al-Hasan, S., Vachtsevanos, G.: Intelligent Route Planning for Fast Autonomous Vehicles Operating in a Large Natural Terrain. *Robotics and Autonomous Systems* 40(1), 1–24
3. Nikolos, I.K., Valavanis, K.P., Tsourveloudis, N.C., et al.: Evolutionary Algorithm Based Offline/Online Path Planner for UAV Navigation. *IEEE Transactions on Systems, Man and Cybernetics, Part B* 33(6), 898–912 (2003)
4. Glmore., J.F., Czuchry, A.J.: A Neural Network Model for Route Planning Constraint Integration. In: *Proceedings of the IEEE 1992 Neural Networks International Joint Conference*, pp. 842–856 (1992)
5. Duan, H.-B., Zhang, X.-Y., Wu, J., et al.: Max-Min Adaptive Ant Colony Optimization Approach to Multi-UAVs Coordinated Trajectory Replanning in Dynamic and Uncertain Environments. *Journal of Bionic Engineering* 6(2), 161–173 (2009)
6. Tian, L.: Research on Algorithm of Multi Route Planning Based on GA for Air Vehicles. Master thesis Shi Jia Zhuang, HeBei Normal University (2006)

7. Chen, D., Zhou, D., Feng, Q.: Route Planning for Unmanned Aerial Vehicles Based on Particle Swarm Optimization. *Journal of Projectiles, Rockets, Missiles and Guidance* 27(4), 340–342
8. Ding, Y., Ren, L.: Artificial Immune Systems: Theory and Applications. *Pattern Recognition and Artificial Intelligence* 13(1), 52–59 (2000)
9. Haibin, D.: *Ant Colony Algorithms: Theory and Applications*, pp. 306–309. Science Press, Beijing (2005)
10. David, F., Yap, W., Habibullah, A., Koh, S.P., Tiong, S.K.: An Improved Artificial Immune System based on Antibody Remainder method for Mathematical Function ptimization. In: *Proceedings of 2010 IEEE Student Conference on Research and Development (SCORED 2010)*, Putrajaya, Malaysia, December 13-14 (2010)
11. Mo, H.: *Artificial Immune System*. Science Press, Beijing (2009)
12. de Castro, L.N.: Learning and Optimization Using the Clonal Selection Principle. *IEEE Transactions on Evolutionary Computation* 6(3) (2002)
13. Bian, Z., Zhang, X.: *Pattern Recognition*. Tsinghua University Press, Beijing (2000)
14. Wen, S., Li, M., Zheng, J.: Comparison and Research of Mutation Operators in Multi-Objective Evolutionary Algorithms. *Computer Engineering and Applications* 45(2), 74–78 (2009)

Hyperspectral Image Classification Using Support Vector Machines with an Efficient Principal Component Analysis Scheme

Pinliang Dong¹ and Jianguo Liu²

¹ Department of Geography, University of North Texas,
Denton, TX 76203, USA
pdong@unt.edu

² Department of Mathematics, University of North Texas,
Denton, TX 76203, USA
jgliu@unt.edu

Abstract. Support vector machines (SVM) together with principal component analysis (PCA) have been applied to hyperspectral image classification and mapping with great success. PCA has been proved to be an effective preprocessing tool for dimension reduction and/or feature extraction. After dimension reduction with PCA, the classification and mapping time can be dramatically reduced while retaining good accuracy. However, the computational cost of PCA preprocessing can be as high as that of SVM classification applied to the original unreduced data set. Researchers have studied different algorithms to cut the PCA preprocessing time, while some others totally ignored that cost. We propose a simple PCA preprocessing scheme which can reduce the computational complexity many folds and is particularly suitable for image data of large size. High classification accuracy can be achieved. A numerical example on an Earth Observing-1 (EO-1) Hyperion image is included to demonstrate the viability of this new procedure. In addition, this example clearly shows that the standard PCA preprocessing may require as much time as the SVM classification does.

Keywords: Support vector machines, hyperspectral image classification, dimension reduction, principal component analysis.

1 Introduction

Digital image classification (and mapping) is a process of abstraction and generalization from image data to produce categories of interest depending on the application. Many methods have been proposed for image classification. In particular, support vector machines (SVM) have been applied to hyperspectral image classification with great success. Satisfactory accuracy can be achieved by SVM when applied to hyperspectral images without preprocessing. See, e.g., [5], [9], and [16].

To overcome the difficulty of massively large data volume and high data dimensionality, and/or to improve classification accuracy, preprocessing on the

data set is usually required in practice. A common preprocessing is dimension reduction and/or feature extraction. Principal component analysis (PCA) is one of the most effective methods for that purpose. After dimension reduction with PCA, the classification time can be dramatically reduced while retaining good accuracy. However, the computational cost of PCA preprocessing can be as high as that of SVM classification applied to the original unreduced data set. Researchers have studied different algorithms to cut the PCA preprocessing time (see, e.g., [4], [10], and [14]), while some others totally ignored that cost. A few research work on SVM and PCA for hyperspectral image classification can be found in [6], [7], [8], [13], [15], and [20].

Some existing fast PCA algorithms can reduce the complexity when the number of pixels to be classified is smaller than the data vector dimension but not the other way around (see, e.g., [14]). In hyperspectral image classification and mapping, the total number of image pixels is usually much larger than the data vector dimension (in this case, the vector dimension is the number of spectral bands). We propose a simple PCA preprocessing scheme which can reduce the computational complexity many folds and is particularly suitable for image data of large size. A numerical example on an Earth Observing-1 (EO-1) Hyperion image is included to demonstrate the viability of this new procedure. In addition, this example clearly shows that the standard PCA preprocessing may require as much time as the SVM classification does.

The rest of the paper is organized as follows. In Section 2, we define the hyperspectral image classification and mapping problem. In Section 3, we describe PCA and dimension reduction. Section 4 briefly states SVM classification. The new PCA procedure is proposed in Section 5. A numerical example is presented in Section 6 and some concluding remarks are given in Section 7.

2 Hyperspectral Image Classification and Mapping

Hyperspectral images are obtained by earth orbiting imaging spectrometers of high spatial resolution. For example, NASA's Earth Observing-1 (EO-1) Hyperion has a resolution of 30m by 30m ([12], [17]). That is, each image pixel represents an area of 30m by 30m on earth. On each pixel, there are hundreds of spectral bands. The EO-1 has 242 bands on each pixel.

To classify and map hyperspectral images, a field engineer may be sent to the corresponding area to collect some class information, i.e., which pixel belongs to what category, such as water, bare soil, vegetation, etc.. Since collecting field information can be very time consuming, only a very small portion of the concerned area can be examined by the field engineer. Then based on the field information, all the given images will be classified and mapped pixel by pixel.

To be more specific, suppose there are k images to be classified. Suppose each image contains p pixels and each pixel has d bands. Let $n = k \times p$ which is the total number of pixels to be classified. Let m be the number of pixels that we have the field information about. We wish to use the m pixels with known class information to train a classifier, then apply the classifier to all n pixels to classify

and map all the given images. Mathematically, we have a matrix X of d rows and n columns. Each column represents a pixel. Given class labels on m of the n columns, we wish to put a label on each of the n columns. It is usually the case that $n \gg m$.

3 Principal Component Analysis

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components [21]. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has as high a variance as possible (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (uncorrelated with) the preceding components.

Let X be the data matrix of d -by- n with zero empirical mean. The number of observations is n and the number of variables is d . Suppose $C = XX^T$ is the covariance matrix of X and suppose $C = W\Sigma W^T$ is the eigenvalue decomposition of C , where Σ is a diagonal matrix containing the eigenvalues of C in descending order, the matrix W is d -by- d and the columns of W are the eigenvectors of C . Then the PCA transformation of X is $Y = W^T X$. The i th row of Y is the i th principal component of X . Suppose Y_h contains the first h rows of Y which are the first h principal components of X . Then Y_h can be calculated by $Y_h = W_h^T X$ where W_h contains the first h columns of W . The matrix W can also be calculated by using the singular value decomposition of X .

4 Support Vector Machine

In its simplest (linear) form, a binary SVM is a hyperplane that separates a set of positive observations from negative observations with given labels by maximizing the class margin (see, e.g., [18] and [19]). The problem can be formulated as a convex quadratic programming (QP) problem and its dual form is often solved in practice. In the linearly inseparable cases, a modification (called the soft margin) is proposed to the QP formulation that allows, but penalizes, observations to fall on the wrong side of the decision boundary [2]. The data set containing the observations with given labels is called the training set. After the QP problem is solved (i.e., the SVM is trained), new unlabeled observations can be classified by the trained SVM. Multiclass classification can be done based on binary classification. Detailed descriptions can be found in [3].

5 An Efficient PCA Scheme

To motivate the proposed PCA scheme, we first analyze the computational complexity of the standard PCA. Once again, let X be the data matrix with d rows

and n columns. Each column represents a pixel and m of the columns are labeled. A subset of the m columns, say, l columns, is usually used to train an SVM while the remaining $m - l$ columns are used for testing the accuracy. Cross validation is often performed. Once a satisfactory SVM is trained, it will be used to classify all the n columns, i.e., to classify and map all the given images.

To reduce classification time, people often use the first few, say, h principal components to do the SVM training (using l labeled observations) and classification (applied to all n observations). The relationship between all the dimension numbers can be summarized by

$$h \ll d < l < m \ll n.$$

For standard PCA, suppose singular value decomposition is used. Then the computational complexity will be $O(nd^2)$ to calculate W and $O(ndh)$ to form Y_h . The overall complexity is $O(nd^2)$ since $h \ll d \ll n$. If eigenvalue decomposition is used, then the complexity will be $O(nd^2)$ to form C , $O(d^3)$ to compute W from C , and $O(ndh)$ to get Y_h . The overall complexity will be $O(nd^2)$. The constant multiple for singular value decomposition is usually of order 10 (see [11]). Therefore, eigenvalue decomposition should be used.

For PCA to work well as a preprocessing tool, we need to assume that the data set is approximately, jointly, and normally distributed since principal components are guaranteed to be independent only if the data set is jointly normally distributed. Under this assumption, we make a key observation: suppose \tilde{X} is the matrix of d -by- l that contains the labeled training observations. If l is reasonably large, then the covariance matrix $\tilde{C} = \tilde{X}\tilde{X}^T$ will be a good approximation to $C = XX^T$ in terms of the PCA transformation. Therefore, suppose $\tilde{C} = \tilde{W}\tilde{\Sigma}\tilde{W}^T$ is the eigenvalue decomposition of \tilde{C} . We may use the approximated PCA matrix $\tilde{Y} = \tilde{W}^T X$ instead of $Y = W^T X$. The complexity to calculate \tilde{W} will be $O(ld^2)$ using singular value decomposition or eigenvalue decomposition. The complexity to form $\tilde{Y}_h = \tilde{W}_h^T X$ is $O(ldh)$. The matrix \tilde{Y}_h can be used to do SVM training and classification. That way, the overall complexity will be reduced from $O(nd^2)$ to $O(ld^2)$ and the total computation time can be reduced many folds. The numerical example in the following section demonstrates effectiveness of the approximated PCA matrix.

6 A Numerical Example on an EO-1 Hyperion Image

The study area is located in the relatively flat Ray Roberts Greenbelt near Denton, Texas, USA (Figure 1). The land cover types are divided into five general categories: (1) dense canopy vegetation (bottomland forest and upland forest), (2) sparse canopy vegetation (rangeland, shrub land, and agricultural land), (3) bare soil, (4) urban, and (5) water. Hyperspectral data in the study area were acquired on August 30, 2004 by NASA's Earth Observing-1 (EO-1) Hyperion and radiometrically corrected by NASA. The original 242 bands were reduced to 138 bands by excluding the bands with (1) all pixels being zeros, (2) low signal to noise ratio (especially, < 430 nm and > 2400 nm), (3) overlap of the VNIR

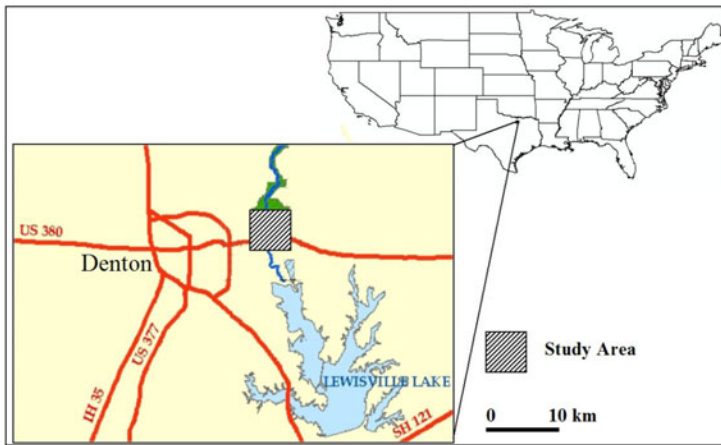


Fig. 1. Study Area

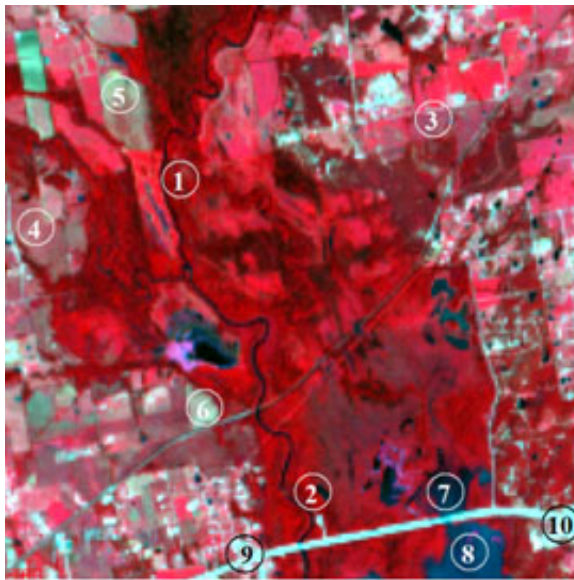


Fig. 2. Color composite of Hyperion image (5.46 km by 5.49 km). The numbers show representative land cover types - 1 and 2: dense canopy vegetation, 3 and 4: sparse canopy vegetation, 5 and 6: bare soil, 7 and 8: water, 9 and 10: urban.

and SWIR spectrometers, and (4) strong atmospheric water absorption. Figure 2 shows the study area with an image of 182 by 183 pixels (5.46 km by 5.49 km). Therefore, the number of images is $k = 1$, the total number of pixels is $n = 182 \times 183 = 33306$. Of the 33306 pixels, $m = 4799$ of them were labeled into

Table 1. CPU time of training, testing, and classification when the unreduced full matrix X is used. The accuracy on the testing set is 92.56%.

Work	CPU Time (seconds)
SVM training	0.12
Testing on the testing set	0.28
Classifying all pixels	2.40
Total	2.80

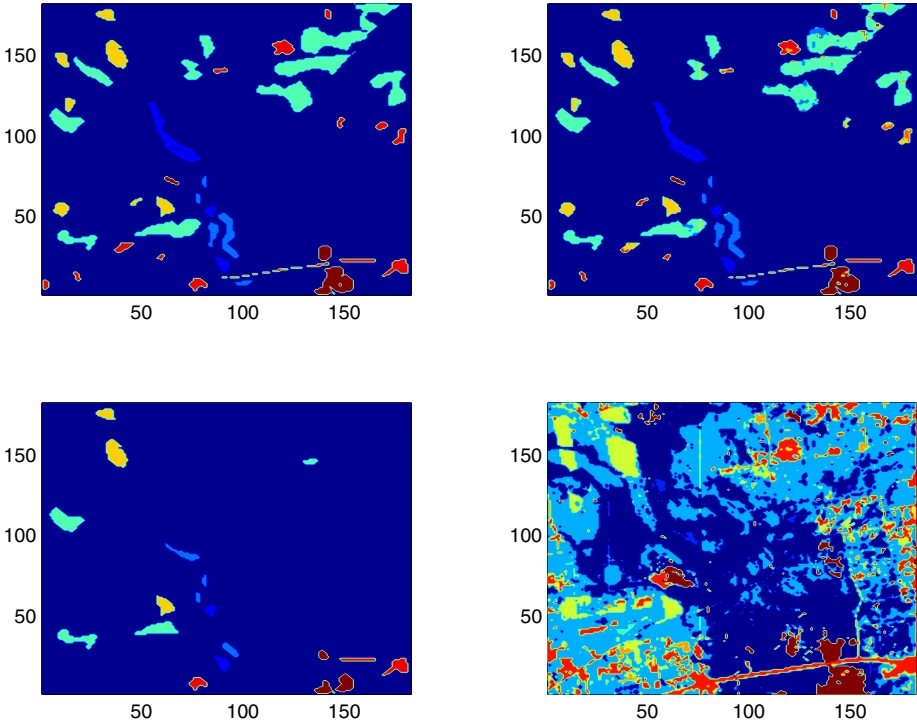


Fig. 3. (1) Upper-left: pixels used for training and testing. (2) Lower-left: pixels used for training. (3) Upper-right: map predicted by the trained SVM on the training and testing pixels. (4) Lower-right: full map generated by the trained SVM on all pixels.

five classes by a field engineer, with $l = 1131$ pixels designated as the training set and the remaining 3368 pixels as the testing and validating set.

For SVM training, testing, and classification, an SVM package LIBSVM [1] was used in a MATLAB environment. The linear polynomial kernel was selected and the regularization parameter C was set to 1. All the computations were performed on a Pentium IV desktop computer with Windows XP operating system.

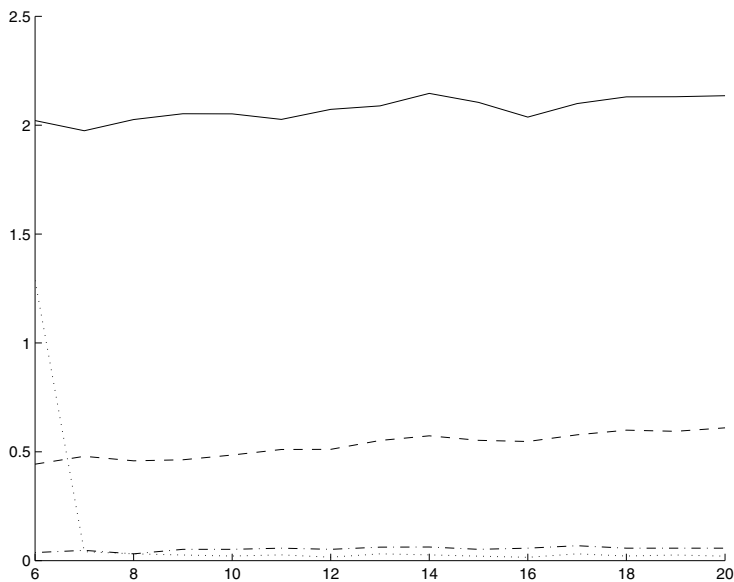


Fig. 4. CPU time (vertical axis in seconds) of the standard PCA preprocessing, training, testing, and classification when the given number of PCs (horizontal axis) are used. (1) solid: PCA; (2) dotted: training; (3) dashdot: testing; (4) dashed: classification.

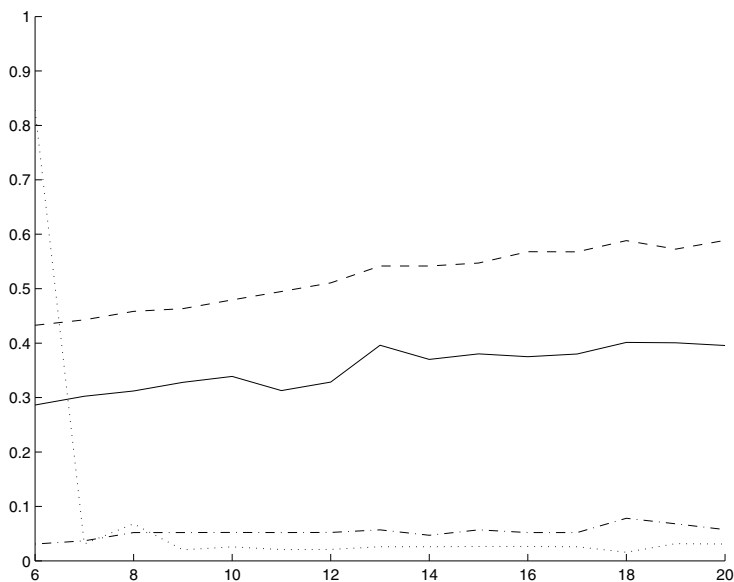


Fig. 5. CPU time (vertical axis in seconds) of the NEW PCA preprocessing, training, testing, and classification when the given number of PCs (horizontal axis) are used. (1) solid-PCA; (2) dotted-training; (3) dashdot-testing; (4) dashed-classification.

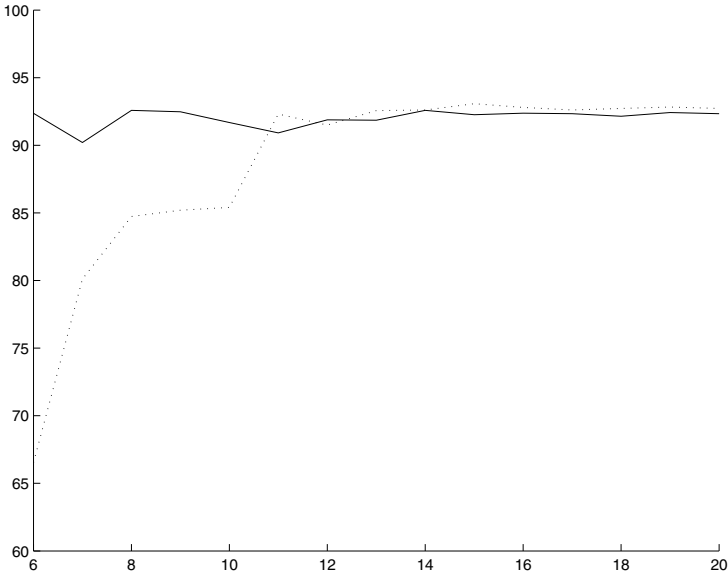


Fig. 6. Accuracy of the trained SVM applied to the testing set (vertical axis in percentage) when the given number of PCs (horizontal axis) are used. (1) solid: using the standard PCA; (2) dotted: using the NEW PCA.

We first ran the training, testing, and classification on the unreduced full matrix X with 138 rows and 33306 columns. The CPU time for each stage is listed in Table 1. For each case, three runs were performed and the number showing is the average. When the trained SVM is applied to the testing set, the accuracy is 92.56%. The training pixels, testing pixels, and the generated map are displayed in Figure 3.

The second run includes the standard PCA preprocessing, followed by SVM training, testing, and classification using different number of principal components (PCs). The average CPU time for each stage with different number of PCs is depicted in Figure 4. For each case, three runs were performed and the number showing is the average. When the number of PCs used is great than 20, accuracy stays the same but CPU time goes up. The accuracy on the testing set is shown in Figure 6. It is clear that the standard PCA preprocessing takes almost as much CPU time as the classification without dimension reduction does. The accuracy is also about the same. Therefore, there is no benefit to use the standard PCA to preprocess the data. We notice that when the number of PCs used is too small, the CPU time for SVM training goes up. The reason is that when too few PCs are used, more support vectors are generated, which results in more calculations in SVM training.

The third run is similar to the second run except the proposed new PCA preprocessing is used instead of the standard PCA, followed by SVM training, testing, and classification using different number of PCs. The average CPU time

for each stage with different number of PCs is depicted in Figure 5. The accuracy on the testing set is also shown in Figure 6.

We see that when the number of PCs used is between 11 and 20, the standard PCA and the proposed new PCA will result in about the same accuracy over 91%. The CPU time on SVM training, testing, and classification will be about the same.

However, for just one image, the CPU time of the PCA preprocessing can be reduced from about 2.1 seconds to 0.3-0.4 seconds. That reduction will be more significant when more images need to be classified and mapped.

7 Concluding Remarks

PCA is a great preprocessing tool for dimension reduction and feature extraction. However, the cost of PCA is sometimes ignored. There are research work on reducing PCA complexity when the number of variables is greater than the number of observations. For hyperspectral image classification and mapping, the number of observation (i.e., the number of pixels) is usually much bigger than the number of variables (i.e., the number of spectral bands). We propose a simple PCA scheme that can reduce the computational complexity many folds. The reduction gets more significant when the number of pixels to be classified increases. The included numerical example demonstrates the efficiency of the new scheme.

Acknowledgments. The authors thank the anonymous referees for their valuable comments.

References

1. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machine (2001), software, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
2. Cortes, C., Vapnik, V.: Support Vector Networks. *Machine Learning* 20, 273–297 (1995)
3. Cristianini, N., Shawe-Taylor, J.: Support Vector Machines and other kernel-based learning methods. Cambridge University Press (2000)
4. El-Bakry, H.M.: A New Implementation of PCA for Fast Face Detection. *World Academy of Science, Engineering and Technology* 16 (2006)
5. Gualtieri, J.A., Cromp, R.F.: Support vector machines for hyperspectral remote sensing classification. In: *Proc. SPIE*, vol. 3584, pp. 221–232 (1998)
6. Jia, X., Richards, J.A.: Segmented Principal Components Transformation for Efficient Hyperspectral Remote-Sensing Image Display and Classification. *IEEE Transactions on Geoscience And Remote Sensing* 37(1) (1999)
7. Kavitha, K., Arivazhagan, S.: Combined Feature Based Hyperspectral Image Classification Technique Using Support Vector Machines. *World Academy of Science, Engineering and Technology* 70 (2010)
8. Lei, H., Govindaraju, V.: Speeding up multi-class SVM evaluation by PCA and feature selection. In: *The 5th SIAM International Conference on Data Mining*, Newport Beach, CA (2005)

9. Mercier, G., Lennon, M.: Support vector machines for hyperspectral image classification with spectral-based kernels. In: International Geoscience and Remote Sensing Symposium, IGARSS, Toulouse, France (September 2003)
10. Neerja ., Walia, E.: Face Recognition Using Improved Fast PCA Algorithm. In: IEEE Congress on Image and Signal Processing (2008)
11. O'Leary, D.P.: Scientific Computing with Case Studies. SIAM (2009)
12. Pearlman, J.S., Barry, P.S., Segal, C.C.: Hyperion, a space-based imaging spectrometer. *IEEE Transactions on Geoscience and Remote Sensing* 41, 1160–1172 (2003)
13. Plaza, A., Benediktsson, J.A., Boardman, J., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, J.A., Marconcini, M., Tilton, J.C., Trianni, G.: Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment* 113(Suppl. 1), S110–S122 (2009)
14. Sharma, A., Paliwal, K.K.: Fast principal component analysis using fixed-point algorithm. *Pattern Recognition Letters* 28, 1151–1155 (2007)
15. Tsai, F., Lin, E.K., Yoshino, K.: Spectrally segmented Principal Component Analysis of hyperspectral imagery for mapping invasive plant species. *Int. J. Remote Sens.* 28, 1023–1039 (2007)
16. Tzotsos, A., Argialas, D.: A support vector machine approach for object based image analysis. In: *Support Vector Machine Classification for Object-Based Image Analysis*, ch. 7.2. Springer (2008)
17. Ungar, S.G., Pearlman, J.S., Mendenhall, J.A.: Overview of the Earth Observing One (EO-1) mission. *IEEE Transactions on Geoscience and Remote Sensing* 41, 1149–1159 (2003)
18. Vapnik, V.: *The Nature of Statistical Learning Theory*. Springer, New York (1995)
19. Vapnik, V.: *Statistical Learning Theory*. John Wiley, New York (1998)
20. Yang, M.D., Huang, K.S., Lin, J.Y., Liu, P.: Application of Support Vector Machines to Airborne Hyper-Spectral Image Classification. In: *Advances in Neural Network Research and Applications*. Lecture Notes in Electrical Engineering, vol. 67(5), pp. 439–444 (2010)
21. Wikipedia: Principal component analysis, http://en.wikipedia.org/wiki/Principal_component_analysis

A Practicable Heuristic Attributes Reduction Algorithm for Ordered Information Systems

Wei Li, Feng Liu, and Zhi-hong Zhao

Supported by the Fundamental Research Funds for the Central Universities(1118021707)
Software Institute, Nanjing University, Nanjing, Jiangsu Province, 210093, China
{mg1032007, liufeng, zhaozh}@software.nju.edu.cn

Abstract. Rough set theory is a mathematic tool to handle vague and uncertain knowledge. The core problem of rough set theory is the reduction of attributes. This paper introduces presentation of the information granularity for ordered information systems based on dominance relation, and detail the way to calculate the attribute importance based on the information granularity we have defined. On that basis, put forward a heuristic attribute reduction algorithm for ordered information systems. Experiments show that the algorithm performs well and outperforms some other rivals.

Keywords: Rough set theory, dominance relationships, attributes reduction, information granularity.

1 Introduction

As a data analysis theory presented by Z.Pawlak in 1982[1], rough set theory is a theoretical tool to deal with fuzziness and uncertainty. It is capable of revealing the potential discipline among imprecise, inconsistent and incomplete information by data analyzing and deducting. Rough set theory has been proven effective during its successful application of many scientific and engineering domains[2-6], such as clinical diagnosis, pattern recognition, machine learning, and image processing etc.

Classic rough set theory uses complete information systems as research objects and bases on the equivalence relation to create equivalence classes from particular domain of objects. However, factors such as noises, information defects and preference information of attributes, increasingly make classic rough set theory inappropriate. Therefore, compatibility relation and similarity relation were brought forward. Greco[7-9] proposed the way based on dominance relation in 1998, which substitutes dominance relation for equivalence relation in the classic rough set and establishes ordered information system to deal with preference problems of standard attributes.

Attributes reduction is a basic problem in rough set theory. Wong.S.K.M and Aiariko.W[10] have proven that the smallest reduction for the decision table of an information system is a NP-hard problem, and one common way to solve it is heuristic searching. Set operation and difference matrixes under dominance relation are widely used. [11,12] discussed the attributes reduction of decision table under dominance relation. [13] introduced attributes reduction of incomplete systems.

Information theory was introduced by a lot of scholars. Different kinds of information entropies were used in rough set theory to describe the uncertainty of knowledge. Miao[14] discussed the relation between roughness of knowledge and information entropies. Wang[15] discussed the relation between algebraic reduction methods and information entropy methods in attributes reduction of decision table. [16] introduced reduction of decision table with relative entropy. Liang[17-18] discussed the relation between knowledge entropy and roughness entropy of rough set in the incomplete systems, and proposed an approach to measure the uncertainty of knowledge. Many practical problems are based on dominance relation, so it is necessary to research the information theory methods of ordered information systems. [19] introduced a way to describe knowledge uncertainty of ordered systems using knowledge entropy and roughness entropy of rough set.

2 Fundamental Theories

Definition 1. Let $I = (U, A, V, f)$ be an information system, where U is a non-empty finite set of objects, A is a non-empty finite set of attributes, $V = \bigcup_{a \in A} V_a$, V_a is the domain of attribute a , for $\forall a \in A$, $x \in U$, $f(x, a) \in V_a$, where $f(x, a)$ is an information function, which means the value of each object x in U .

Every attribute set determines a binary indistinguishable relation in rough set theory that based on the equivalence relation, which has obvious limitations in practical application. By establishing a partial order relation, either by descending or not, we can establish an information system based on dominance relation.

Definition 2. An attribute is a *criterion* when a partial order relation is established in the domain of it in an information system. If all the attributes of the information system are criteria, the system itself is an ordered information system.

Let $I = (U, A, V, f)$ be an information system, $a \in A$ is a criterion. Define a partial order relation \geq_a in the domain of attribute a , where $x \geq_a y$ means x is as good as y for the criterion a , or x is better.

In general, let domain of attributes be real numbers, and define $\geq_a: x \geq_a y \Leftrightarrow f(x, a) \geq f(y, a)$. For attribute set $B \subseteq A$, $x \geq_B y$ means x is better than y for all the criteria in the attribute set B .

Definition 3. Given an information system $I = (U, A, V, f)$, for attribute set $B \subseteq A$, $R_B^{\geq} = \{(x, y) \in U \times U \mid f(x, a) \geq_a f(y, a), \forall a \in B\}$, where R_B^{\geq} is defined as the dominance relation of ordered information system $I = (U, A, V, f)$.

Let $[x]_B^{\geq} = \{y \in U \mid f(y, a) \geq_a f(x, a), \forall a \in B\}$, $U / R_B^{\geq} = \{[x]_B^{\geq} \mid x \in U\}$, where $[x]_B^{\geq}$ is defined as the dominance class, U / R_B^{\geq} as the partition for all the objects in the domain refer to attribute set B .

Property 1. Let $I = (U, A, V, f)$ be an information system, assume $B \subseteq A$

- 1) R_B^{\geq} is reflexive and transitive, but not always symmetrical, so it is not an equivalence relation any more in general;
- 2) given that $P \subseteq Q \subseteq A$, then $R_A^{\geq} \subseteq R_Q^{\geq} \subseteq R_P^{\geq}$, $[x]_A^{\geq} \subseteq [x]_Q^{\geq} \subseteq [x]_P^{\geq}$;
- 3) if $x \in [x]_B^{\geq}$, then $[y]_B^{\geq} \subseteq [x]_B^{\geq}$;
- 4) only if $f(x, a) = f(y, a)$, $[y]_B^{\geq} = [x]_B^{\geq}$;

5) U / R_B^{\geq} covers all the objects in U , namely $\forall x \in U, [x]_B^{\geq} \neq \emptyset$ and $\cup_{x \in U} [x]_B^{\geq} = U$.

Definition 4. $\forall X \subseteq U$, define the lower approximation of X referring to dominance relation R_B^{\geq} as: $R_B^{\geq}(X) = \{x \in U | [x]_B^{\geq} \subseteq X\}$; the upper approximation of X as: $\overline{R_B^{\geq}}(X) = \{x \in U | [x]_B^{\geq} \cap X \neq \emptyset\}$.

Definition 5. Given an ordered information system (U, A, V, f) , where $B, B' \subseteq A$, define the partial relation: $B \preceq B' \Leftrightarrow \forall x \in U, [x]_B^{\geq} \subseteq [x]_{B'}^{\geq}$. Donote $B \preceq B'$ if B' is coarser than B (or B is finer than B' as the same); if $B \preceq B'$ and $B \neq B'$, we say B' is strictly coarser than B (or B is strictly finer than B' as the same), denote as $B < B'$.

Definition 6. Given an ordered information system $I = (U, A, V, f)$, $b \in B \subseteq A$. If $U / R_B^{\geq} = U / R_{B-\{b\}}^{\geq}$, we say b is not necessary in B , otherwise b is necessary in B . If each $b \in B$ is necessary, we say B is independent, or B is dependent if not.

Definition 7. Let $I = (U, A, V, f)$ be an information system, if $B \subseteq A$ and

1) $R_B^{\geq} = R_A^{\geq}$, 2) B is independent,

B is a reduction of ordered information system I , and define the denotation: $\subseteq Red(A)$, where $Red(A)$ is the set of all the reductions of ordered information system I .

Definition 8. Let $I = (U, A, V, f)$ be an information system, we define the core of A referring to dominance relation R_A^{\geq} , and denotation of it $core(A)$, which is the set of all necessary attributes in attribute set A .

Property 2. $core(A) = \cap Red(A)$, and the core may be null.

3 Information Granularity and Attribute Importance

In this section, we introduce an information granularity model to describe the amount of information of the attribute set, and the denotation of attribute importance will be defined in the following as well.

Definition 9. Let $I = (U, A, V, f)$ be an information system, $\subseteq A$, define the information granularity of B :

$$GI(B) = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|[x_i]_B^{\geq}|}{|U|}, x_i \in U. \tag{1}$$

Property 3. Let $I = (U, A, V, f)$ be an information system, $\subseteq A$, then $GI(B)$ gets minimum $\frac{1}{|U|}$ when $U / R_B^{\geq} = \{[x_i]_B^{\geq} = \{x_i\} | x_i \in U\}$, and maximum 1 when $U / R_B^{\geq} = \{[x_i]_B^{\geq} = U | x_i \in U\}$.

Theorem 1. Let $I = (U, A, V, f)$ be an information system, $B, B' \subseteq A$, if $B = B'$, $GI(B) = GI(B')$.

Proof. Since $B = B'$, we have $\forall x_i \in U, [x_i]_B^{\geq} = [x_i]_{B'}^{\geq}$, hence, $\frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|[x_i]_B^{\geq}|}{|U|} = \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|[x_i]_{B'}^{\geq}|}{|U|}$, so it follows immediately that $GI(B) = GI(B')$.

Theorem 2. Let $I = (U, A, V, f)$ be an information system, $B, B' \subseteq A$, if $B \preccurlyeq B'$, $GI(B) \leq GI(B')$.

Proof. Since $B \preccurlyeq B'$, we have $\forall x_i \in U, [x_i]_B^{\geq} \subseteq [x_i]_{B'}^{\geq}$, namely $|[x_i]_B^{\geq}| \leq |[x_i]_{B'}^{\geq}|$, hence, $\frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|[x_i]_B^{\geq}|}{|U|} \leq \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|[x_i]_{B'}^{\geq}|}{|U|}$, and it comes out that $GI(B) \leq GI(B')$.

Deduction 1. Let $I = (U, A, V, f)$ be an information system, and $P, Q \subseteq A$. If $P \subseteq Q$, $GI(P) \geq GI(Q)$.

Proof. Since $P \subseteq Q$, we have $R_Q^{\geq} \subseteq R_P^{\geq}$, and $\forall x_i \in U, [x_i]_Q^{\geq} \subseteq [x_i]_P^{\geq}$, namely $|[x_i]_Q^{\geq}| \leq |[x_i]_P^{\geq}|$. Hence, we have $\frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|[x_i]_Q^{\geq}|}{|U|} \geq \frac{1}{|U|} \sum_{i=1}^{|U|} \frac{|[x_i]_P^{\geq}|}{|U|}$, and finally $GI(P) \geq GI(Q)$.

Definition 10. Let $I = (U, A, V, f)$ be an information system, for $\forall a \in B, B \subseteq A$, attribute importance

$$sig_{in}(a, B) = GI(B - \{a\}) - GI(B) . \tag{2}$$

Property 4. Let $I = (U, A, V, f)$ be an information system, for $\forall a \in B, B \subseteq A$, the information importance $sig_{in}(a, B)$ has following properties:

- 1) $0 \leq sig_{in}(a, B) \leq 1 - \frac{1}{|U|}$;
- 2) a is necessary if and only if $sig_{in}(a, B) > 0$;
- 3) $(A) = \{a | sig_{in}(a, A) > 0, a \in A\}$.

Definition 11. Let $I = (U, A, V, f)$ be an information system, for $\forall a \in A - B, B \subseteq A$, attribute importance

$$sig_{out}(a, B) = GI(B) - GI(BU\{a\}) . \tag{3}$$

Based on the nearest two definitions above, we draw the conclusion that the importance of an attribute can be described as the increase of information granularity if we remove it from the attribute set; and for a particular attribute set, the importance of a new attribute to be added can be described as the decrease of information granularity if we add it into the set.

Theorem 3. Let $I = (U, A, V, f)$ be an information system, the sufficient condition for attribute set $B \subseteq A$ to be a reduction of I is:

- 1) $GI(B) = GI(A)$;
- 2) for $\forall a \in B$, $sig_{in}(a, B) > 0$.

Proof. We can know from (2) that B is independent; since $(B) = GI(A)$, we say the information granularities of A and B are equal referring to Theorem1, and for $\forall a \in B$, $sig_{in}(a, A) > 0$.

4 An Attribute Reduction Algorithm

We will use the attribute importance model introduced above to show our attribute reduction algorithm in this section. Suppose $I = (U, A, V, f)$ is an ordered information system, we will calculate the similar set of core attribute set of I , as the reduction of it. At first, we calculate attribute importance of every attribute in A with the attribute importance model, and range them in descending order. Then we pick up those attributes in order according to their importance from high to low and calculate the importance relative to the current reduction. The calculation won't stop until the decrease of the information granularity of a certain attribute's addition to the reduction is smaller than the predefined threshold α , or the difference between the information granularity of the current reduction set and I is smaller than the predefined threshold β . The reduction set we've got is the reduction set of information set I .

Algorithm 1. An attribute reduction algorithm of ordered information systems.

Input: an ordered information system $I = (U, A, V, f)$;

Output: attribute reduction (A) ;

Algorithm steps:

Step1. Let $Red(A) = \emptyset$;

Step2. for ($i=1; i \leq |A|; i++$) {

Calculate $sig_{in}(a_i, A)$;

Insert a_i into another ordered array A_{Seq} according to its importance;

}

Step3. Let $sig_{min} = 0$;

for ($j=1; j \leq |A_{Seq}|; j++$) {

calculate $sig_{in}(a_j, Red(A))$;

$sig_{min} = \text{Min}\{sig_{in}(a_j, Red(A))\}$;

if $sig_{in}(a_j, Red(A)) \leq \alpha$ and $GI(Red(A)) - GI(A) \leq \beta$

break;

else

$Red(A) = Red(A) \cup \{a_j\}$;

}

Time complexity for calculating importance of every attribute a_i in step2 is $O(|U|^2)$, and time complexity for sorting is $O(|A|^2)$. Considering that $|A|$ is far more smaller than $|U|$, time complexity of this step is then $O(|U|^2|A|)$. Time complexity for calculating importance of every attribute a_j in step3 is $O(|U|^2)$, and $O(|U|^2) + O(|Red(A)|^2)$ for calculating $GI(Red(A)) - GI(A)$, so time complexity for this step should be $(|U|^2|A|)$. Therefore, the total time complexity of the attribute reduction algorithm is $(|U|^2|A|)$.

5 Experiments

Based on the attribute reduction algorithm described above, we use the *Sendmail* program experiment data publicly released by University of New Mexico in America to evaluate the performance and reliability of our algorithm, and also do some comparison with other reduction algorithm.

The experiment is divided into training step and verifying step. The data set we used contains 7 normal sequences, 5 fwd-loops wrong sequences, 3 sscp attacking sequences, 4 syslog attacking sequences, 2 decode attacking sequences and 2 unsuccessful intrusion sequences. We choose some of them for training and then the left for verifying.

Each of the sequences is a system call set produced by a process, and it records all the system calls in order of their invocation during the whole execution of the process. In case of the invocation of fork, each kind of processes is allowed to have several process IDs. We denote the system call flow of a process as $Seq_{ab} = (c_1, c_2, \dots, c_l)$, which belongs to category a and has the ID b , c_i represents the position and the number of system calls. We record the probability of different system calls that appear in the flow, and range them. Suppose there are m different system calls in the flow, we can use an array named as *probability* of length m to store the appearance probability of each call in descending order. Therefore, $probability[i]$ represents the appearance probability of the system call whose position in the array is i . The consecutive system call flow can't be used directly, so we have to cut them into several segments with the same length. We use a sliding window of fixed length k to extract those system calls by the size of window from the first to the end of the flow Seq_{ab} , and each window moves a fixed number of steps forward from the former one, which is 1 in this experiment. In this way, we get $l - k + 1$ short sequences of size k , and the one of position j can be represented as $S_j = (c_j, c_{j+1}, \dots, c_{j+k-1})$. To go a step further, we replace the position of appearance probability of each system call rather than the call itself in each sequences, and we finally get $S'_j = (c'_j, c'_{j+1}, \dots, c'_{j+k-1})$, where c'_j represents the position of the appearance probability of system call c_j in the array *probability*.

We denote all the processes as unique IDs with positive integers, and mark the processes according to their categories. Suppose that a process of category a is denoted as t_a with S_j as the conditional attribute set and $\{d\}$ as the decision attribute set, then we define the decision item $S'_j = (c'_j, c'_{j+1}, \dots, c'_{j+k-1}, \{d\})$, which means if a short sequence like $S_j = (c_j, c_{j+1}, \dots, c_{j+k-1})$ appears in the process of category t_a in the training, the process may belong to category t_a or the others as well.

Accordingly, we establish an ordered information system $I = (U, A, V, f)$ based on dominance relation, where $|U| = l - k + 1$, $|A| = k$. We use *Algorithm 1* to perform attribute reduction for ordered information system I , and the result will be named as *rule*.

System call flow of a process to be verified is also be divided into several short sequences in the way we introduce above in the training step. We compare each short sequence with the rule and record that how much they are matched. If a short sequence matches the rule well, it can match at least one of the categories that the processes in the rule belong to.

We use array $match_j[]$ to represent whether each short sequence S_j matches different process categories. The index t of the array represents the category of the process. Initially, for every t , $match_j[t] = 0$, which means S_j hasn't be matched to category t . We try to match S_j to every item in the rule one by one. If it matches to category t successfully once, value of $match_j[t]$ is updated to 1. We use an array $count[]$ to store the number of short sequences matched to each category. Suppose number of short sequences is n , then $count[t] = \sum_{j=1}^n match_j[t]$, $0 \leq t < T$. The counter variable $unmatch$ is used to represent number of short sequences that can't be match to any category. If S_j doesn't match to any of the item in the rule, $unmatch$ will be increased by 1.

We define support and confidence as the basis to verify whether the short sequences of some kind of process is correctively recognized. Support is denoted as $Sup = 1 - unmatch/n$, confidence $Conf[t] = (count[t])/(n - unmatch)$. Support Sup represents the percentage of short sequences successfully matched in the total ones. If the quota is too low, the process being tested may be unsuccessfully matched to all of the categories in the training set. Confidence $Conf[t]$ represents the percentage of the short sequences matched to its category in the total ones. If the confidence gets very low, it means the process being tested may not be well matched to the current category.

Define a threshold V_{sup} for support Sup , and V_{conf} for confidence $nf[t]$. If $Sup < V_{sup}$, the process being tested will be marked as unknown. Otherwise, we find the t that makes $Conf[t]$ maximum. If $nf[t] < V_{conf}$, the current process will also be marked as unknown, or category t instead.

6 Analysis of Experiment

We do experiments with the window size 15, 20, 25, 30 for comparison.

According to the experiment, for the same α and β , the bigger the window size is, the longer both of the time of reduction and verification are. For different α and β , reduction time increases at first and then decreases as the window size increases, and the support fluctuates while the differences between the confidences of processes that be matched and the others increase. All of the results above are under expectation. The bigger the window size is, the more comparing work is in the reduction and verification. With the increase of α and β , number of attributes that satisfy the condition $sig_{in}(a, A) > \alpha$ and $sig_{out}(a, B) > \beta$ increases at first and then decreases, so time for reduction also changes identically. Because knowledge of each system call won't change and increased window size also increases the knowledge of each short sequence, information granularity of each attribute set decreases. Therefore, rule extracted from the process of particular category has more complete information of it, and the category of the process itself can be distinguished more easily as well. This explains why the differences between the confidences of processes that be matched and the others increase. However, support reflects the recognition probability of short sequences, so it fluctuates as a result of the variety categories of process.

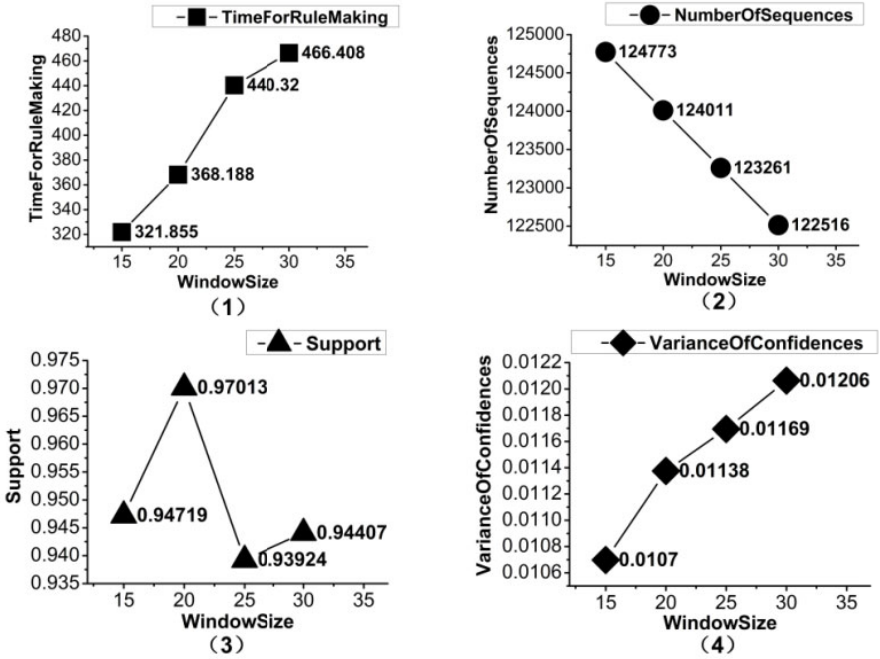


Fig. 1. Given that $\alpha = 0.05, \beta = 0.4$. (1) Time for attribute reduction; (2) number of short sequences; (3) support; (4) variance of confidence

We also do some comparisons with other reduction algorithms. With the same window size, our algorithm's performance is better than Qian[20], and much better than the value reduction algorithm proposed by Gu[21] in evidence. However, it is not that better than value reduction algorithm in respect of effect, so does the capability of classification.

7 Conclusions

With the introduction of dominance relation and information granularity, we propose a practicable heuristic attributes reduction algorithm based on the conception of attribute importance. This algorithm can create a reduction of an ordered information system with relatively reliable capability of classification. It gains great increase while compared with difference matrix algorithms in performance as well. The experiments show that the algorithm can be used to verify unknown system call sequences, thus it may give a way to establish a decision support system.

References

1. Pawlak, Z.: Roughsets. *International Journal of Computer and Information Science* 11, 341–356 (1982)
2. Chmielewski, M.R., Grzymala-Busse, J.W.: Global discretization of continuous attributes as preprocessing for machine learning. *International Journal of Approximate Reasoning* 15, 319–331 (1996)
3. Lingras, P.J., Yao, Y.Y.: Data mining using extensions of the rough set model. *Journal of the American Society for Information Science* 49(5), 415–422 (1998)
4. McSherry, D.: Knowledge discovery by inspection. *Decision Support Systems* 21, 43–47 (1997)
5. Pomerol, J.C.: Artificial Intelligence and Human Decision Making. *European Journal of Operational Research* 99, 3–25 (1997)
6. Slowinski, R.: *Intelligent Decision Support: Handbook of Applications and Advances of the Rough Sets Theory*, pp. 23–33. Kluwer Academic Publishers, Dordrecht (1992)
7. Greco, S., Matarazzo, B., Slowinski, R.: Rough approximation of a preference relation by dominance relation. *European Journal of Operation Research* 117, 63–83 (1999)
8. Greco, S., Matarazzo, B., Slowinski, R.: A New Rough Set Approach to Multicriteria And Multiattribute Classification. In: Polkowski, L., Skowron, A. (eds.) *RSCTC 1998. LNCS (LNAI)*, vol. 1424, pp. 60–67. Springer, Heidelberg (1998)
9. Greco, S., Matarazzo, B., Slowinski, R.: Rough approximation by dominance relations. *International Journal of Intelligent System* 17(2), 153–171 (2002)
10. Wong, S.K.M., Ziarko, W.: On Optimal Decision Rules in Decision Tables. *Bulletin of Polish Academy of Sciences* 33(11-12), 693–696 (1985)
11. Zhang, W.X., Liang, Y., Wu, W.Z.: *Information Systems and Knowledge Discovery*. Science Press, Beijing (2003) (in Chinese)
12. Xu, W.H., Zhang, W.Y.: Knowledge Reductions in Inconsistent Information Systems Based on Dominance Relations. *Computer Science* 33(2), 182–184 (2006)
13. Shao, M.W., Zhang, W.X.: Dominance Relation and Rules in an Incomplete Ordered Information System. *International Journal of Intelligent Systems* 20, 13–27 (2005)
14. Miao, D.Q., Wang, Y.: An Information Representation of the Concepts and Operations in Rough Set Theory. *Journal of Software* 10(2), 113–116 (1999)
15. Wang, G.Y.: Algebra View And Information View of Rough Sets Theory: Data Mining and Knowledge Discovery Theory, Tools, and Technology III. In: *Proceeding of SPIE*, vol. 4384, pp. 200–207 (2001)
16. Gui, X.C.: Attribute Reduction Algorithm Based on the Relative Entropy. *Computer Engineering and Applications* 42(33), 157–159 (2006)
17. Liang, J.Y., Xu, Z.B.: The Algorithm on Knowledge Reduction in Incomplete Information Systems. *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems* 24(1), 95–103 (2002)
18. Liang, J.Y., Li, D.Y.: *Uncertainty in Information Systems and Knowledge Acquisition*. Science Press, Beijing (2005) (in Chinese)
19. Zhang, X.Y., Xu, W.H.: Entropy of Knowledge and Rough Set in Ordered Information Systems. *Computer Engineering and Applications* 43(27), 62–65 (2007)
20. Wang, Y., Qian, Y.H., Liang, J.Y.: Heuristic Attribute Reducton Algorithm to Ordered Information Systems. *Computer Science* 1, 258–260 (2010)
21. Gu, J.H., Zhou, Y.C., Song, J., Yan, J.Q.: A New Attribute Value Reduction Algorithm. *Acta Scientiarum Naturalium University Nankaiensis* 36(4), 38–42 (2003) (in Chinese)

An Improved Abrams-Strogatz Model Based Protocol for Agent Competition and Strategy Designing

Cunhua Li, Yun Hu, and Lanlan Sun

School of Computer Engineering, Huaihai Institute of Technology
Lianyungang, Jiangsu Province, P.R. China. 222005
cli@hhit.edu.cn

Abstract. Strategy designing is a key task in today's market place competition and many other application fields. While dominating over competitors is unrealistic or too expensive to seek after, finding dynamic equilibrium in the competition progress becomes critical and practical in lots of scenarios. Inspired by the Abrams and Strogatz model originally proposed for monitoring language death and competition, this paper presented a novel mechanism for agent's competition equilibrium point finding and strategy designing. By analyzing historical data and status property of the two competition agents, an agent could evaluate its own status and forecast the long range trends of the game based upon the improved Abrams and Strogatz model. A posteriori status evaluation based strategy designing protocol is purposed and implemented. Experiments and simulation results showed desirable consistence with theoretical analysis. The protocol presented is valuable for adoption with proper settings to practical applications.

Keywords: Abrams Strogatz model, agent competition, strategy design, equilibrium.

1 Introduction

Competition is the key character of the nature life, the evolution of the society and every sector the market place[1]. However, It is very common in today's world that the counterparts in a competition environment try to seek a compromise for the purpose of resource saving or to avoid vicious competition. This is true in many aspect of today's social life, such as arms race among countries or competition for customers between enterprises in the market [1,2]. For example, a service provider will try to keep his own share of customers in the market at a relatively steady percent other than to monopolize it because of the law enforcement or too expensive to do that.

Competition strategy is widely researched in game theory and agent coalition formation [2,3]. In game theory, the most important aspect of strategy design is to understand the opponent's position, and correctly deducing the likely response for its own. The Nash equilibrium is a set of strategies for players to do its best, given the strategies of the other players. In the research of competitive agent coalition formation, the center problem is to find agent subset in a competitive environment which can most effectively carry out a particular task where coalitions compete amongst themselves

for a payoff [3]. Research area includes problems such as finding strategies for agent attraction, churn prevention and coalition stability, finding stability properties between coalition structures [4].

In linguistics research area, researchers model the competition among languages just as biological species. The initial model proposed by Abrams and Strogatz (the AS Model) [5] treat the language as an agent competing for customers. The models focus on the evolution of the number of speaker in the population in terms of first order differential rate equations, ignoring internal structure of the language such as syntax, grammar and their changes.

In this paper, we utilize the dynamics of the AS model to study how an agent can keep equilibrating against its competitor in a continuous step by step manner. A set of equilibrium seeking and strategy formation mechanism is proposed depending on the basis of self status evaluation and competitor's strategy deduction. We optimize the parameter of the model to fit the model for real world problems such as agent in the retail market. A Posteriori Status Evaluation based Strategy Designing protocol (PoStESD Protocol) is purposed for the agent.

This paper is organized as follows: In Section 2 we introduce the Abrams and Strogatz models and briefly review earlier results on language competition. In Section 3, we discuss our results of simulation on the stable and unstable equilibrium point of the AS model. How to keep the two competitor agents under relative balance near the unstable equilibrium point of the system is also studied in this section. This section also gives a strategy designing protocol for the agent to keep track of its counterpart's strategy while ensure the balance of the competition. In Section 4 we perform a set of stability analysis of the balance keeping competition with particularly selected initial state of the system. These results demonstrate that unstable equilibrium is retainable for long range if only the agent have the ability of status management by choosing correct strategies in each step.

2 The Abrams Strogatz Model

In 2003, Abrams and Strogatz proposed a model for language competition [5]. The model analytically represented the competition between languages in a quantitative and dynamic manner. After then, many researchers had revised the Abrams and Strogatz model by considering the existence of bilingualisms and the connectivity of people in language competition [6-8].

Let X and Y be two languages competing for speakers. The percentage of population speaking each language is denoted as x and y respectively. Each individual only speak one of the two language, such that $x + y = 1$. An individual switches from one language (for example X) to its competing partner (Y) is due to the "attractiveness" of the competing language. The attractiveness of a language increases with both its number of speakers and its perceived status, denoted as s . The parameter s can be understood as some factor of advantage that a particular language offered to its speakers. The relative status for language X and Y are denoted by s_X and s_Y , and $s_X + s_Y = 1$.

Suppose an individual speaker converts from Y to X with a probability $P_{YX}(x, s_X)$ per unit time. Likewise, the convert probability from X to Y can be written

as $P_{XY}(y, s_Y)$. With all of the above assumption and notation, the rate at which the fraction of population speaking X changes can be introduced based on the following first order differential rate equation:

$$\frac{dx}{dt} = yP_{YX}(y, s_X) - xP_{XY}(x, s_Y). \tag{1}$$

The differential equation for Y can be wrote the same way because the symmetric property between X and Y.

For the probability $P_{YX}(x, s_X)$, Abrams and Strogatz further proposed its functional form driven by the power-law as follows.

$$P_{YX}(y, s_X) = cs_Xx^\alpha \tag{2}$$

thus,

$$P_{XY}(x, s_Y) = cs_Yy^\alpha = c(1-s_x)(1-x)^\alpha \tag{3}$$

Here α is a parameter that models how the attractiveness of a language scales with the number of its speakers. Based on fitting with historical data, α is roughly a constant across cultures, with $\alpha = 1.31 \pm 0.25$.

Plugging the expression for $P_{YX}(x, s_X)$ from equation (2) and (3) into equation (1), it follows,

$$\frac{dx}{dt} = c(1-x)s_xx^\alpha - cx(1-s_x)(1-x)^\alpha. \tag{4}$$

The equation (4) governs the future of language on whether it will be booming or go extinction depending on the initial conditions. The two stable equilibrium states of the equation, $x = 0$ and $x = 1$, are trivial because they means X will go extinction or dominance in the long run. What we indeed interested in (4) is its unstable equilibrium state, $0 < x < 1$. For each given x , we can always find a s_x that satisfies $dx/dt = 0$. It confirmed that it is possible to have multiple competing languages coexist by controlling their status based on the fraction of population speaking each language.

The idea of AS model inspired us to look into many similar problems in the real world. There are many competition circumstances in which “one aspect wanes, the other waxes”. Keeping one’s own strength in the competition is the first goal for survival! When characterizing real world problems, the AS model is valuable in monitoring the trend of competition and enable for reasonable strategy choosing to adjust the future result of the competition. This is the basis of our following discussion.

3 The Posteriori Status Evaluation Strategy Designing Protocol

Although the Abrams and Strogatz model is originally proposed to monitor competition between languages, however, the principle of the model is worthy reference to monitor a lot of scenarios in which competition exists. As an example, this paper generalize the AS model to study the transition mechanism between two parties who

share the same market of valuable service. We proposed a strategy designing algorithm based on Posteriori Status Evaluation (PSE) of competition status and the last step action of the counterpart. Modifications on the AS model are done in two aspect. While the AS model is strictly determined by the set of initial parameters, we suppose that the process is adjustable by modifying the parameters in step-by-step manner. Thus, we can monitor the competition process in each step and act as a character in the game by strategy making to change the route of the long range evolution. This is what we called the Posteriori Status Evaluation based strategy designing. Another optimization we made is on accumulative effects of the parameters. While each step of original AS model is only determined by its nearest predecessor state, we argue that it is reasonable to consider the accumulative effects of the former states.

Suppose X and Y are two broker agents sharing and serving the whole market of N ($N \gg 1$) customers. Each customer chooses only one of the two brokers as its own service provider at each time period. The customer may switch from one broker (for example X) to its competitor at a random rate depend on the attractiveness function s_x of the broker X on the accumulative basis of its k cycle of time period. That is,

$$\tilde{s}_x = \sum_{i=0}^{-k} \alpha_i s_x^i \tag{5}$$

Here, s_x^0 is the immediate predecessor attractiveness value of agent X and s_x^{-i} is the attractiveness value of X in the i th former state, $\alpha_i, i=0, -1, \dots, -k$ the weight of importance for attractiveness evaluation. Just like in Abrams Strogatz model, we denote by x and y as the proportion of the market share of X and Y, and $x + y = 1$.

Let $P_{YX}(x, s_x)$ be the probability function of customers changing from Y to X per unit of time. Then, the rate of customers belonging to X is characterized by the following equation:

$$\frac{dx}{dt} = yP_{YX}(y, \tilde{s}_x) - xP_{XY}(x, \tilde{s}_y). \tag{6}$$

Let $x^{(k)}, y^{(k)}, s_x^{(k)}, s_y^{(k)}$ be the proportion and status of broker agent X and Y respectively in step k. And $P_{XY}^{(k)}(y, s_y)$ be the combination set of strategies broker Y have chosen in the step k. Then, with these input parameters, the numerical solution of the equation (6) can be simply deduced by the Euler iteration,

$$x^{(k+1)} = x^{(k)} + (y^{(k)}P_{YX}(y^{(k)}, \tilde{s}_x^{(k)}) - x^{(k)}P_{XY}(x^{(k)}, \tilde{s}_y^{(k)})) \cdot \Delta t. \tag{7}$$

At this point, broker X is ready to take two kind of advantage of the Euler iteration (6) for forecasting of its own fortune. The one is an evaluation of its attractiveness change in the market simply use the variation between $x^{(k+1)}$ and $x^{(k)}$. The other is a long range simulation research using the AS model to show the fate of its own and the counterpart Y. The attractiveness equilibrium point cab be computed or by running a simulation. The expected value of $s_x^{(k+1)}$ to the equilibrium state is determined by solving equation $dx/dt \geq 0$. that is,

$$y^{(k+1)}P_{YX}(y^{(k+1)}, s_x^{(k+1)}) - x^{(k+1)}P_{XY}(x^{(k+1)}, s_y^{(k+1)}) \geq 0 \tag{8}$$

The mechanism of our key idea in posteriori status evaluation based strategy designing protocol can be described using the following flow chart:

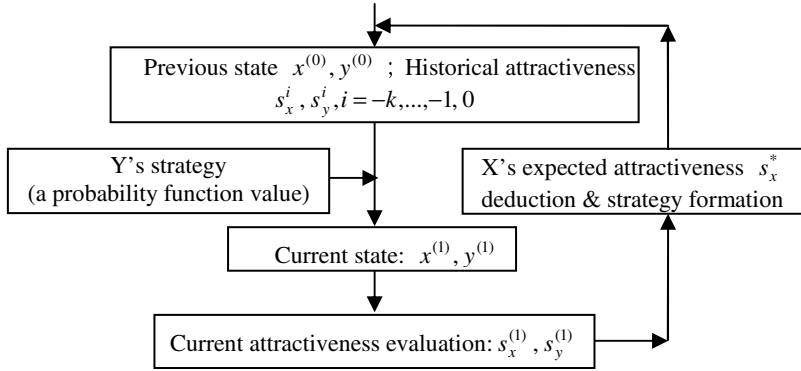


Fig. 1. The flow chart for X's status analysis and strategy designing

The above Abrams Strogatz model-based strategy designing process starts with the record of both party's previous states: $x^{(0)}, y^{(0)}$, historical attractiveness of X and Y: $s_x^i, s_y^i, i = 0, -1, \dots, -k$. Before the system coming into the current state, counterpart Y may choose some kind of strategy to boost its attractiveness. In our simulation case, we randomly generate a probability function value for Y because Y is not controllable by X. With the above set of parameters, we now can obtain the current status of X and Y: $x^{(1)}, y^{(1)}$ using formula (7). An evaluation of both party's attractiveness $s_x^{(1)}, s_y^{(1)}$ can be processed at the same time. The expected equilibrium point s_x^* for X against its competitor Y can be solved using formula (8) or by running methodical simulation with the AS model. A comparison between X's actual attractiveness value $s_x^{(1)}$ and the expected value s_x^* indicates whether the current status of X is healthy enough to withstand the competition from Y. If $s_x^{(1)} < s_x^*$, the AS model tells that X's condition will derogate in the future and means for improving its attractiveness must be taken.

4 Experiment Results

To examine the dynamics of the strategies explained and validate the theoretical analysis, a range of simulation experiments were conducted. Each simulation run consists of a set of iterations with varied parameters. In the equilibrium state keeping oriented competition simulation experiments, our agent X follow the iterated PoStESD protocol explained in the above section for solving a number of different tasks with differential initial state.

In the first set of experiments, we used the original AS model to solve the threshold s_x^* for the equilibrium state of the competition game, given X with different initial proportion of shares. The results showed that the AS model is very sensitive to both of the initial population and initial status.

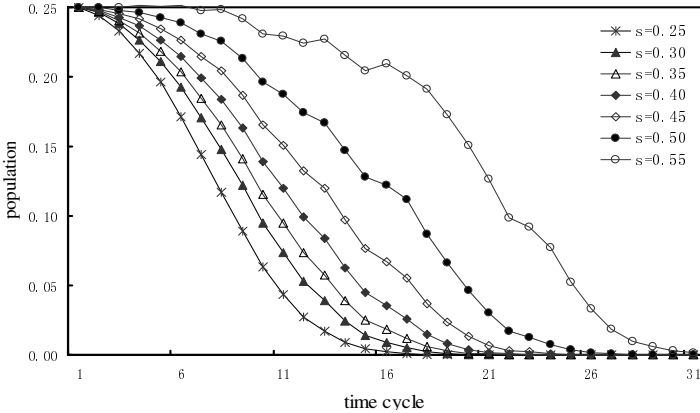


Fig. 2. Simulation results using Abrams Strogatz model start with initial population $x=0.25$ and various attractiveness values s from 0.25 to 0.55. For $s \geq 0.585$, X will win over Y finally

Fig.2 shows how fast agent X will die out with an initial proportion $x = 0.25$ and different attractiveness status s . when $s \geq 0.585$, X will win over Y if the model runs automatically. Given any initial value of x and s_x , by adjusting the status parameters of the model steadily, we can always keep the system revolving around its equilibrium state as long as the adjustment goes on.

To realize a long range coexistence and compromise state between X and Y , we designed a program for agent X using the PoStESD protocol for its expected attractiveness values solving (Fig. 3).

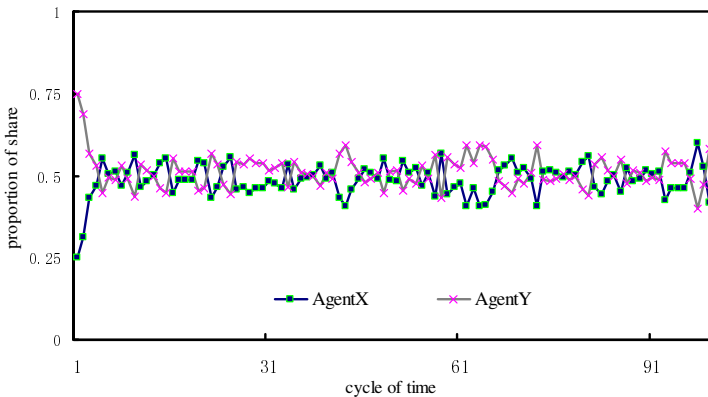


Fig. 3. A scenario of long period (100 cycle) coexistence of X and Y

In our experiments, the proportion of market share of X and Y was 0.25 versus 0.75. To X, this was really trying to fight the giant as a tiny nobody! Under the AS model, it is easy to see that X have no chance to be alive for a long time because it has smaller population and lower attractiveness status. However, if we could intervene the process by adjusting the attractiveness value of X when needed, the way of competition may change gradually.

For visual and explanation convenient, we first chose larger value for s_x in the early stage of the experiments so that X could gain half of the population share. After the compromise state achieved, X's task was to monitor the variation of both party's population shares and choose correct value of next s_x such that the system would keep the equilibrium state with a proper range of perturbation. Fig. 3 illustrates the simulation results of first 100 cycles of the equilibrium point keeping oriented competition.

5 Conclusions and Further Work

Market and many forms of competition have become data-driven in nowadays world. By analyzing data collected from the market using proper models, one can be confident to what is going on the next step or the long range trends of its own, and, in some degree, of its competitors. Choosing corrected models is critical and always of great interest to researchers and enterprises. This paper introduced an Abrams Strogatz model based tools for status analysis and competition simulation, especially for circumstances where equilibrium is the main consideration. The Abrams Strogatz model based methods have solid mathematical background and powerful ability in modeling a series of real world problem. Our current work is on revised Abrams Strogatz model based problems which involves multi-party and multi-factor competition.

Acknowledgments. We acknowledge financial support of the the Natural Science Foundation of Jiangsu Province under Grant No. BK2008190.

References

1. Spyros, G.M.: Forecasting, Planning, and Strategy for the 21st Century. The Free Press, A Division of Macmillan, Inc., New york (1990)
2. Ken-Ichi, S., Jacques-François, T.: Competition Among the Big and the Small. CREA Discussion Paper Series 09-18, Center for Research in Economic Analysis, University of Luxembourg (2009), <http://ideas.repec.org/p/luc/wpaper/09-18.html>
3. McAfee, R.P., John, M.: Competition and Game Theory. *Journal of Marketing Research* 33(8), 263–267 (1996)
4. Merida-Campos, C., Willmott, S.: Agent Compatibility and Coalition Formation: Investigating Two Interacting Negotiation Strategies. In: Fasli, M., Shehory, O. (eds.) TADA/AMEC 2006. LNCS (LNAI), vol. 4452, pp. 75–89. Springer, Heidelberg (2007)

5. Abrams, D.M., Strogatz, S.H.: Modeling the dynamics of language death. *Nature* 424, 900 (2003)
6. Kandler, A., Steele, J.: Ecological models of language competition. *Biology Theory* (3), 164–173 (2008)
7. Kosmidis, K., Halley, J.M., Argyrakis, P.: Language evolution and population dynamics in a system of two interacting species. *Physica A* (353), 595–612 (2005)
8. Stauffer, D., Schulze, C.: Microscopic and macroscopic simulation of competition between languages. *Life Rev.* (2), 89–116 (2005)

α -Quasi-Lock Semantic Resolution Method for Linguistic Truth-Valued Lattice-Valued Propositional Logic $\mathcal{L}_{V(n \times 2)}P(X)$

Xiaomei Zhong¹, Jun Liu², Shuwei Chen², and Yang Xu¹

¹ School of Mathematics, Southwest Jiaotong University,
Chengdu, Sichuan 610031, P.R. China

zhongxm@126.com, xuyang@home.swjtu.edu.cn

² School of Computing and Mathematics, University of Ulster at Jordanstown,
Newtownabbey, BT37 0QB, Northern Ireland, UK

j.liu@ulster.ac.uk, chen-s1@email.ulster.ac.uk,

Abstract. On the basis of α -quasi-lock semantic resolution method in lattice-valued propositional logic $(\mathcal{L}_n \times \mathcal{L}_2)P(X)$, α -quasi-lock semantic resolution in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ is studied in the present paper. Firstly, (c_i, t) -quasi-lock semantic resolution for $\mathcal{L}_{V(n \times 2)}P(X)$ is equivalently transformed into that for lattice-valued propositional logic $\mathcal{L}_{V_n}P(X)$. Secondly, similar equivalence between (c_i, f) -quasi-lock semantic resolution for $\mathcal{L}_{V(n \times 2)}P(X)$ and that for $\mathcal{L}_{V_n}P(X)$ is also established under certain conditions.

Keywords: α -Quasi-lock semantic resolution method, Resolution-based automated reasoning, Linguistic truth-valued lattice-valued propositional logic, Linguistic truth-valued lattice implication algebra.

1 Introduction

In order to make machines can simulate some intelligent behavior of human, such as reasoning, judgements and decision-making, it is very necessary to study the automated reasoning based on many-valued logic. Lattice-valued logic system based on lattice implication algebras is a kind of important many-valued logic, which can describe both comparability and incomparability simultaneously. Since a gradational resolution principle-- α -resolution principle [10] was proposed in this lattice-valued logic, resolution-based automated reasoning in this lattice-valued logic, which can be used for program verification, has become an important issue in the corresponding research field.

Considering people in real life usually use linguistic truth values to describe truth values of actual problems, and in many cases, people also use linguistic truth values to make reasoning, judgment and decision-making, so in order to use linguistic truth values to implement α -resolution automated reasoning based on the above lattice-valued logic, Xu, et al., established linguistic truth-valued lattice implication algebra $\mathcal{L}_{V(n \times 2)}$ [12] for modeling some linguistic truth values frequently used in daily life.

Therefore the research about resolution-based automated reasoning methods in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ based on linguistic truth-valued lattice implication algebra $\mathcal{L}_{V(n \times 2)}$ has become extremely important. As α -quasi-lock semantic resolution method based on the above lattice-valued logic can improve the efficiency of resolution in two ways, this paper will further study the α -quasi-lock semantic resolution method based on linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$.

In order to study resolution-based automated reasoning for lattice-valued propositional logic LP(X) based on lattice implication algebras, in 2000, Xu et al. [10], proposed the α -resolution principle based on LP(X), and its soundness and weak completeness theorems were also proved. In 2006, Xu et al. [12], established linguistic truth-valued lattice implication algebra $\mathcal{L}_{V(n \times 2)}$ by some linguistic truth values frequently used in daily life. In 2008, Li et al. [5], studied properties of the α -resolution field and J -resolution field in lattice-valued propositional logic $\mathcal{L}_{n \times 2}P(X)$, and also discussed the relation between α -resolution and J -resolution in $\mathcal{L}_{n \times 2}P(X)$. In 2010, He et al. [9], proposed α -lock resolution method based on lattice-valued propositional logic $\mathcal{L}_n P(X)$ and established its soundness and weak completeness. At the same year, to further improve the efficiency of α -resolution, Xu et al. [13], proposed the general form of α -resolution principle in the above lattice-valued logic, and proved its soundness and weak completeness theorems. Based on the general form of α -resolution principle, we established α -quasi-lock semantic resolution method [7] based on LP(X), as well as its soundness and weak completeness in 2011.

This paper is organized as follows: Section 2 reviews some preliminary relevant concepts and conclusions involved in this paper. In Section 3, (c_i, t) -quasi-lock semantic resolution for $\mathcal{L}_{V(n \times 2)}P(X)$ is studied. In Section 4, (c_i, f) -quasi-lock semantic resolution for $\mathcal{L}_{V(n \times 2)}P(X)$ is studied.

2 Preliminaries

In the following, we will introduce some elementary concepts and conclusions involved in this paper. We refer readers to [7, 11] for more details.

2.1 Lattice Implication Algebra

Definition 2.1. [11] Let (L, \vee, \wedge, O, I) be a bounded lattice with an order-reversing involution $'$, I and O the greatest and the smallest element of L respectively, and $\rightarrow: L \times L \rightarrow L$ be a mapping. $(L, \vee, \wedge, ', \rightarrow, O, I)$ is called a lattice implication algebra if the following conditions hold for any $x, y, z \in L$:

$$\begin{aligned} (I_1) \quad & x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z), & (I_2) \quad & x \rightarrow x = I, & (I_3) \quad & x \rightarrow y = y' \rightarrow x', \\ (I_4) \quad & x \rightarrow y = y \rightarrow x = I \text{ implies } x = y, & (I_5) \quad & (x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x, \\ (I_6) \quad & (x \vee y) \rightarrow z = (x \rightarrow z) \wedge (y \rightarrow z), & (I_7) \quad & (x \wedge y) \rightarrow z = (x \rightarrow z) \vee (y \rightarrow z). \end{aligned}$$

Example 2.1. [11] (**Łukasiewicz implication algebra on finite chain**) Let $L_n = \{a_i \mid i = 1, 2, \dots, n\}$, $a_1 < a_2 < \dots < a_n$. For any $1 \leq j, k \leq n$, define

$$a_j \vee a_k = a_{\max\{j, k\}}, \quad a_j \wedge a_k = a_{\min\{j, k\}}, \quad (a_j)' = a_{n-j+1}, \quad a_j \rightarrow a_k = a_{\min\{n-j+k, n\}},$$

then $(L_n, \vee, \wedge, ', \rightarrow, a_1, a_n)$ is a lattice implication algebra, denoted by \mathcal{L}_n .

Example 2.2. [12] Let $\mathcal{L}_n = (L_n, \vee_1, \wedge_1, ', \rightarrow_1, a_1, a_n)$ be the Łukasiewicz implication algebra in Example 2.1. $L_2 = \{b_1, b_2\}$, $b_1 < b_2$, $\mathcal{L}_2 = (L_2, \vee_2, \wedge_2, ', \rightarrow_2, b_1, b_2)$ is also a Łukasiewicz implication algebra. For any $(a_i, b_j), (a_k, b_m) \in L_n \times L_2$, define

$$(a_i, b_j) \vee (a_k, b_m) = (a_i \vee_1 a_k, b_j \vee_2 b_m), \quad (a_i, b_j) \wedge (a_k, b_m) = (a_i \wedge_1 a_k, b_j \wedge_2 b_m),$$

$$(a_i, b_j)' = (a_i', b_j'), \quad (a_i, b_j) \rightarrow (a_k, b_m) = (a_i \rightarrow_1 a_k, b_j \rightarrow_2 b_m),$$

then $(L_n \times L_2, \vee, \wedge, ', \rightarrow, (a_1, b_1), (a_n, b_2))$ is a lattice implication algebra, denoted by $\mathcal{L}_n \times \mathcal{L}_2$.

Definition 2.2. [12] Let $AD_n = \{c_1, c_2, \dots, c_n\}$ be a set with n modifiers and $c_1 < c_2 < \dots < c_n$, $MT = \{f, t\}$ be a set of meta truth values, and $f < t$. Denote $L_{V(n \times 2)} = AD_n \times MT$.

Define a mapping g as $g: L_{V(n \times 2)} \rightarrow \mathcal{L}_n \times \mathcal{L}_2$,

$$g((c_i, mt)) = \begin{cases} (a_i', b_1), & mt = f, \\ (a_i, b_2), & mt = t, \end{cases}$$

then g is a bijection, denote its inverse mapping as g^{-1} . For any $x, y \in L_{V(n \times 2)}$, define

$$x \vee y = g^{-1}(g(x) \vee g(y)), \quad x \wedge y = g^{-1}(g(x) \wedge g(y)), \quad x' = g^{-1}((g(x))'), \quad x \rightarrow y = g^{-1}(g(x) \rightarrow g(y)),$$

then $\mathcal{L}_{V(n \times 2)} = (L_{V(n \times 2)}, \vee, \wedge, ', \rightarrow, (c_n, f), (c_n, t))$ is called a linguistic truth-valued lattice implication algebra generated by AD_n and MT .

Remark 2.1. In the following, Łukasiewicz implication algebras $(AD_n, \vee, \wedge, ', \rightarrow, c_1, c_n)$ and $(MT, \vee, \wedge, ', \rightarrow, f, t)$ are denoted by $\mathcal{L}_{Vn}, \mathcal{L}_{V2}$ respectively, where $\vee, \wedge, '$ and \rightarrow are the same as those in Example 2.1.

Example 2.3. [12] Let $AD_9 = \{\textit{Slightly (Sl for short), Somewhat (So), Rather (Ra), Almost (Al), Exactly (Ex), Quite (Qu), Very (Ve), Highly (Hi), Absolutely (Ab)}\}$ be the set of modifiers, $MT = \{\textit{False (f for short), True (t)}\}$ the set of meta truth values. According to Definition 2.2, we can obtain the linguistic truth-valued lattice implication algebra $\mathcal{L}_{V(9 \times 2)}$ generated by AD_9 and MT , where $\mathcal{L}_{V(9 \times 2)} = (L_{V(9 \times 2)}, \vee, \wedge, ', \rightarrow, (Ab, f), (Ab, t))$, $L_{V(9 \times 2)} = AD_9 \times MT$.

2.2 Lattice-Valued Propositional Logic LP(X)

Definition 2.3. [11] Let X be the set of propositional variables, $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $T = L \cup \{', \rightarrow\}$ be a type with $\text{ar}(') = 1$, $\text{ar}(\rightarrow) = 2$ and $\text{ar}(a) = 0$ for any $a \in L$. The proposition algebra of the lattice-valued proposition

calculus on the set X of propositional variables is the free T algebra on X and denoted by $LP(X)$.

Definition 2.4. [11] The set \mathcal{F} of formulas of lattice-valued propositional logic $LP(X)$ is the least set Y satisfying the following conditions:

- (1) $X \subseteq Y$, (2) $L \subseteq Y$, (3) if $p, q \in Y$, then $\neg(p), \rightarrow(p, q) \in Y$,

where X is the set of propositional variables, L is the set of constants.

In the following, we denote $\neg(p)$ as p' and $\rightarrow(p, q)$ as $p \rightarrow q$.

Definition 2.5. [11] A mapping $v: LP(X) \rightarrow L$ is called a valuation of lattice-valued propositional logic $LP(X)$, if it is a T -homomorphism.

Definition 2.6. [11] Let $G \in \mathcal{F}$ and $\alpha \in L$. If $v(G) \leq \alpha$ for any valuation v of lattice-valued propositional logic $LP(X)$, we say G is α -false, denoted by $G \leq \alpha$ (or $G = \alpha \odot$).

Definition 2.7. [11] A lattice-valued propositional logical formula G in $LP(X)$ is called an extremely simple form, in short ESF, if a logical formula G^* obtained by deleting any constant or literal or implication term occurring in G is not equivalent to G .

Definition 2.8. [11] A lattice-valued propositional logical formula G in $LP(X)$ is called an indecomposable extremely simple form, in short IESF, if the following two conditions hold:

- (1) G is an ESF containing connectives \rightarrow and $'$ at most,

(2) for any $H \in \mathcal{F}$, if $\overline{H} \in \overline{G}$ in $LP(X)$, then H is an ESF containing connectives \rightarrow and $'$ at most, where $\overline{LP(X)} = (\overline{\mathcal{F}}, \vee, \wedge, ', \rightarrow)$ is a lattice implication algebra,

$\overline{\mathcal{F}} = \mathcal{F}/\equiv = \{ \overline{p} \mid p \in \mathcal{F} \}$, $\overline{p} = \{ q \mid \text{for any valuation } v \text{ in } LP(X), v(q) = v(p) \}$, for any $\overline{p}, \overline{q} \in \overline{\mathcal{F}}$, $\overline{p} \vee \overline{q} = \overline{p \vee q}$, $\overline{p} \wedge \overline{q} = \overline{p \wedge q}$, $(\overline{p})' = \overline{p'}$, $\overline{p} \rightarrow \overline{q} = \overline{p \rightarrow q}$.

All constants, literals and IESFs in $LP(X)$ are called generalized literals. The disjunction of a finite number of generalized literals is a generalized clause.

Theorem 2.1. [6] Let g_1, g_2, \dots, g_m be generalized literals in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ based on linguistic truth-valued lattice implication algebra $\mathcal{L}_{V(n \times 2)}$, $\alpha = (c_i, d) \in L_{V(n \times 2)}$, then $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq \alpha$ if and only if one of the following two conditions holds:

(1) $d = t$, $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq c_i$, where g_1, g_2, \dots, g_m are interpreted in $\mathcal{L}_{V_n}P(X)$;

(2) $d = f$, $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq c_{n+i}$ and $g_1 \wedge g_2 \wedge \dots \wedge g_m = f$, where g_1, g_2, \dots, g_m are interpreted in $\mathcal{L}_{V_n}P(X)$ and $\mathcal{L}_{V_2}P(X)$ respectively.

Definition 2.9. [9] Let C be a generalized clause in $LP(X)$. C is called a locked generalized clause if each disjunct occurring in C is assigned a positive integer in its lower

left corner (the same disjunct appearing in different locations can be labeled different positive integer). The positive integer is called a lock of the disjunct.

Definition 2.10. [13] Let $C_i = p_{i1} \vee \dots \vee p_{im_i}$ be generalized clauses in $LP(X)$, $H_i = \{p_{i1}, \dots, p_{im_i}\}$ the set of all disjuncts occurring in C_i , $i = 1, 2, \dots, m$, $\alpha \in L$. For any $i \in \{1, 2, \dots, m\}$, if there exist $x_i \in H_i$ such that $x_1 \wedge x_2 \wedge \dots \wedge x_m \leq \alpha$, then $C_1(x_1 = \alpha) \vee C_2(x_2 = \alpha) \vee \dots \vee C_m(x_m = \alpha)$ is called an α -resolvent of C_1, C_2, \dots, C_m , denoted by $R_{p(g-\alpha)}(C_1(x_1), C_2(x_2), \dots, C_m(x_m))$, x_1, x_2, \dots, x_m are called an α -resolution group.

Definition 2.11. [7] Let v_0 be a valuation in lattice-valued propositional logic $LP(X)$, $\alpha \in L$. N, E_1, \dots, E_q are sets composed of some locked generalized clauses in $LP(X)$. The sequence (N, E_1, \dots, E_q) is called an α -quasi-lock semantic clash (α -QLS clash for short) w.r.t. v_0 , if N, E_1, \dots, E_q satisfy the following conditions:

- (1) $v_0(C_i) \leq \alpha$, for any generalized clause $C_i \in E_i$, $i = 1, 2, \dots, q$,
- (2) let $R_0 = \bigvee_{C \in N} C$. For any $i = 1, 2, \dots, q$, there exists an α -resolvent R_i of N_i and E_i , where $N_1 = N$ and for any $i = 2, \dots, q$, $N_i = R_{i-1} \cup N_i^*$, $N_i^* \subseteq N \cup \{R_1, \dots, R_{i-2}\}$,
- (3) for any generalized clause $C_i \in E_i$, the α -resolution literal g_i of C_i is the one that has the smallest lock among disjuncts occurring in C_i , $i = 1, 2, \dots, q$,
- (4) for any generalized clause $C_j \in N_j$, the α -resolution literal g_j of C_j is the one which is non- α -false under valuation v_0 and has the smallest lock among non- α -false disjuncts (under v_0) occurring in C_j , where $j = 1, 2, \dots, q$,
- (5) $v_0(R_q) \leq \alpha$,

R_q is called the α -QLS resolvent (w.r.t. v_0) of this clash. E_1, \dots, E_q are called electrons and N is called the core of this clash.

3 α -Quasi-Lock Semantic Resolution Method in $\mathcal{L}_{V(n \times 2)}P(X)$

As all elements occurring in linguistic truth-valued lattice implication algebra $\mathcal{L}_{V(n \times 2)}$ can be divided into two parts, i.e., (c_i, t) and (c_j, f) , so we will discuss α -quasi-lock semantic resolution in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ based on $\mathcal{L}_{V(n \times 2)}$ by two cases.

In the following, let v_1, v_2 be valuations in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ and $\mathcal{L}_{V_n}P(X)$ whose truth-value field is \mathcal{L}_{V_n} , respectively. Since propositional variables and constant symbols can be interpreted in any logic system, so we can take the following convention.

Convention 3.1. Let (c_i, d_j) , p be constant symbol and propositional variable in $\mathcal{L}_{V(n \times 2)}P(X)$, respectively, we take the following convention:

- (1) if $v_1(p)$ (or $v_1((c_i, d_j))$) = (c_j, t) , then $v_2(p)$ (or $v_2((c_i, d_j))$) = c_j ;
(2) if $v_1(p)$ (or $v_1((c_i, d_j))$) = (c_j, f) , then $v_2(p)$ (or $v_2((c_i, d_j))$) = $c'_j = c_{n-j+1}$.

3.1 (c_i, t) -Quasi-Lock Semantic Resolution Method in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$

Theorem 3.1. Let g be a generalized literal in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$, v_1 a valuation in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$ and $\alpha = (c_i, t) \in L_{V(n \times 2)}$, then the following two results hold:

(1) $v_1(g) \not\leq \alpha$ if and only if $v_2(g) > c_i$, (2) $v_1(g) \leq \alpha$ if and only if $v_2(g) \leq c_i$, where v_2 is the valuation in $\mathcal{L}_{V_n}\mathbf{P}(X)$ that satisfies Convention 3.1.

Proof. Case (1): (\Rightarrow) Let $v_1(g) = (c_j, t)$ or (c_k, f) , $c_j, c_k \in AD_n$. Since $v_1(g) \not\leq \alpha$, so $c_j > c_i$ and $c_k < c_{n-i+1}$. According to Convention 3.1, we can obtain the following results:

- ① if $v_1(g) = (c_j, t)$, then $v_2(g) = c_j > c_i$;
② if $v_1(g) = (c_k, f)$, then $v_2(g) = c'_k > c_i$ by $c_k < c_{n-i+1}$.

Hence, $v_2(g) > c_i$.

(\Leftarrow) Let $v_2(g) = c_j$, so we can obtain $v_1(g) = (c_j, t)$ or (c'_j, f) by Convention 3.1. As $v_2(g) > c_i$, so $c'_j < c'_i$. Hence, we have $(c_j, t) > \alpha$ and $(c'_j, f) \parallel \alpha$, i.e., $v_1(g) \not\leq \alpha$.

Case (2): According to Case (1), the result holds obviously.

Corollary 3.1. Let C be a generalized clause in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$, v_1 a valuation in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$ and $\alpha = (c_i, t) \in L_{V(n \times 2)}$, then the following two results hold:

(1) $v_1(C) \not\leq \alpha$ if and only if $v_2(C) > c_i$, (2) $v_1(C) \leq \alpha$ if and only if $v_2(C) \leq c_i$, where v_2 is the valuation in $\mathcal{L}_{V_n}\mathbf{P}(X)$ that satisfies Convention 3.1.

Proof. According to Theorem 3.1, we can obtain these results easily.

Theorem 3.2. Let $S = C_1 \wedge C_2 \wedge \dots \wedge C_m$, where C_1, C_2, \dots, C_m are locked generalized clauses in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$. v_1 is a valuation in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$ and $\alpha = (c_i, t) \in L_{V(n \times 2)}$. Then the following two statements are equivalent:

- (1) there exists an α -QLS resolution deduction (w.r.t. v_1) from S to $\alpha \odot$,
(2) there exists an c_i -QLS resolution deduction (w.r.t. v_2) from S to $c_i \odot$, where v_2 is the valuation in $\mathcal{L}_{V_n}\mathbf{P}(X)$ that satisfies Convention 3.1.

Proof. Case 1: Suppose D_1 is an α -QLS resolution deduction (w.r.t. v_1) from S to $\alpha \odot$. Then for any α -QLS clash (N, E_1, \dots, E_q) of D_1 , we can obtain (N, E_1, \dots, E_q) is a c_i -QLS clash (w.r.t. v_2). In fact, for any generalized clause $C \in N, G \in E_1 \cup \dots \cup E_q$, we have $v_2(C) > \alpha$ and $v_2(G) \leq \alpha$ by Corollary 3.1. For any α -resolution group (g_1, g_2, \dots, g_k) occurring in α -QLS clash (N, E_1, \dots, E_q) , (g_1, g_2, \dots, g_k) is also an c_i -

resolution group by Theorem 2.1. According to Theorem 3.1, we can draw the conclusion that (N, E_1, \dots, E_q) is a c_i -QLS clash (w.r.t. v_2). So replacing all α in D_1 with c_i , we obtain a c_i -QLS resolution deduction (w.r.t. v_2) D_2 from S to c_i - \odot .

Case 2: The proof is similar to that of Case 1.

Example 3.1. Let $C_1 = x \rightarrow y$, $C_2 = (x \rightarrow z)' \vee (s \rightarrow u)$, $C_3 = y' \vee (y \rightarrow z) \vee (s \rightarrow (Al, t))$, $C_4 = (s \rightarrow u)' \vee (r \rightarrow (So, f))'$, $C_5 = r \rightarrow (Ex, t)$ be generalized clauses in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(9 \times 2)}P(X)$ based on $\mathcal{L}_{V(9 \times 2)}$ and $S = C_1 \wedge C_2 \wedge C_3 \wedge C_4 \wedge C_5$, where $\mathcal{L}_{V(9 \times 2)}$ is the lattice implication algebra in Example 2.3, (Al, t) , (So, f) , $(Ex, t) \in L_{V(9 \times 2)}$ and x, y, z, r, s, u are propositional variables. If $\alpha = (Qu, t)$, then there exists an α -QLS resolution deduction from S to α - \odot .

According to Theorem 3.2, we only need to prove that there exists a Qu -QLS resolution deduction from S to Qu - \odot when S is interpreted in lattice-valued propositional logic $\mathcal{L}_{V_9}P(X)$ based on lattice implication algebra \mathcal{L}_{V_9} . Let S have the following lock:

- $C_1 = {}_1(x \rightarrow y)$,
- $C_2 = {}_2(x \rightarrow z)' \vee {}_3(s \rightarrow u)$,
- $C_3 = {}_4y' \vee {}_5(y \rightarrow z) \vee {}_6(s \rightarrow (Al, t))$,
- $C_4 = {}_7(s \rightarrow u)' \vee {}_8(r \rightarrow (So, f))'$,
- $C_5 = {}_9(r \rightarrow (Ex, t))$.

Suppose v_2 is the valuation in $\mathcal{L}_{V_9}P(X)$ such that $v_2(x) = Ab$, $v_2(y) = v_2(z) = So$, $v_2(s) = Qu$, $v_2(t) = Ra$, $v_2(r) = Qu$. Hence we have $v_2(C_1) < Qu$, $v_2(C_4) < Qu$, $v_2(C_2) > Qu$, $v_2(C_3) > Qu$, $v_2(C_5) > Qu$, and so there exists the following Qu -QLS resolution deduction ω^* :

- (1) ${}_1(x \rightarrow y)$
- (2) ${}_2(x \rightarrow z)' \vee {}_3(s \rightarrow u)$
- (3) ${}_4y' \vee {}_5(y \rightarrow z) \vee {}_6(s \rightarrow (Al, t))$
- (4) ${}_7(s \rightarrow u)' \vee {}_8(r \rightarrow (So, f))'$
- (5) ${}_9(r \rightarrow (Ex, t))$
- (6) ${}_6(s \rightarrow (Al, t)) \vee {}_3(s \rightarrow u) \vee Qu$ by (1), (2), (3)
- (7) ${}_3(s \rightarrow u) \vee {}_8(r \rightarrow (So, f))' \vee Qu$ by (4), (6)
- (8) ${}_8(r \rightarrow (So, f))' \vee Qu$ by (4), (5), (7)
- (9) Qu by (5), (8).

Hence there exists a Qu -QLS resolution deduction (w.r.t. v_2) from S to Qu - \odot . Furthermore, we can obtain an α -QLS resolution deduction ω (w.r.t. v_1) from S to α - \odot by replacing all Qu occurring in ω^* with α , where v_1 is the valuation in $\mathcal{L}_{V(9 \times 2)}P(X)$ that satisfies Convention 3.1.

According to Theorem 3.2, we can find the fact that (c_i, t) -QLS resolution for linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ and c_i -QLS resolution for lattice-valued propositional logic $\mathcal{L}_{V_n}P(X)$ are equivalent.

3.2 (c_i, f) -Quasi-Lock Semantic Resolution Method in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$

In the following, let g be a generalized literal and C a generalized clause in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$, then g^Δ and C^Δ satisfy the following two conditions:

- (1) if there is no constant occurring in g (or C), then $g^\Delta = g$ (or $C^\Delta = C$),
- (2) if there exist constants occurring in g (or C), then g^Δ (or C^Δ) is obtained by replacing (c_r, f) and (c_k, t) with c_1, c_n respectively, for any constants $(c_r, f), (c_k, t)$ occurring in g (or C).

Theorem 3.3. Let g_1, g_2, \dots, g_m be generalized literals in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$, $\alpha = (c_i, f) \in L_{V(n \times 2)}$ and $c_i > c_1$. For any $j = 1, 2, \dots, m$, if $g_j^\Delta \leq g_j$ (g_j^Δ, g_j are interpreted in $\mathcal{L}_{V_n}\mathbf{P}(X)$), then $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq \alpha$ if and only if $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq c_{n-i+1}$.

Proof. (\Rightarrow) Suppose $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq \alpha$, so $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq c_{n-i+1}$ by Theorem 2.1.

(\Leftarrow) Since for any $j = 1, 2, \dots, m$, $g_j^\Delta \leq g_j$ and $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq c_{n-i+1}$, so we have $g_1^\Delta \wedge g_2^\Delta \wedge \dots \wedge g_m^\Delta \leq c_{n-i+1} < c_n$. Suppose $V = \{v \mid v \text{ is a valuation in } \mathcal{L}_{V_n}\mathbf{P}(X) \text{ and } v(p) = c_1 \text{ or } c_n, \text{ for any propositional variable } p \text{ occurring in } g_j, j = 1, 2, \dots, m\}$. Therefore, for any valuation $v_0 \in V$, $v_0(g_1^\Delta \wedge g_2^\Delta \wedge \dots \wedge g_m^\Delta) = c_1$. As L_2 only has the greatest element b_2 and the smallest elements b_1 , hence $g_1 \wedge g_2 \wedge \dots \wedge g_m = b_1$, where g_1, g_2, \dots, g_m are interpreted in $\mathcal{L}_{V_2}\mathbf{P}(X)$. Hence, we can obtain $g_1 \wedge g_2 \wedge \dots \wedge g_m \leq \alpha$ by Theorem 2.1.

Theorem 3.4. Let g be a generalized literal in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$, v_1 a valuation in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$ and $\alpha = (c_i, f) \in L_{V(n \times 2)}$, then the following two results hold:

- (1) $v_1(g) \geq (c_{i-1}, f)$ if and only if $v_2(g) \geq c_{n-i+2}$,
 - (2) if $v_1(g) \leq \alpha$, then $v_2(g) \leq c_{n-i+1}$,
- where v_2 is the valuation in $\mathcal{L}_{V_n}\mathbf{P}(X)$ that satisfies Convention 3.1.

Proof. Similar to the proof of Theorem 3.1, we can obtain the results.

Corollary 3.2. Let C be a generalized clause in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$, v_1 a valuation in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$ and $\alpha = (c_i, f) \in L_{V(n \times 2)}$, then the following two results hold:

- (1) $v_1(C) \geq (c_{i-1}, f)$ if and only if $v_2(C) \geq c_{n-i+2}$,
 - (2) if $v_1(C) \leq \alpha$, then $v_2(C) \leq c_{n-i+1}$,
- where v_2 is the valuation in $\mathcal{L}_{V_n}\mathbf{P}(X)$ that satisfies Convention 3.1.

Proof. According to Theorem 3.4, the results hold obviously.

Theorem 3.5. Let $S = C_1 \wedge C_2 \wedge \dots \wedge C_m$, where C_1, C_2, \dots, C_m are locked generalized clauses in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$. v_1 is a

valuation in $\mathcal{L}_{V(n \times 2)}\mathbf{P}(X)$, $\alpha = (c_i, f) \in L_{V(n \times 2)}$ and $c_i > c_1$. If the following conditions hold:

- (1) for any disjunct g in S , if g, g^Δ are interpreted in $\mathcal{L}_{V_n}\mathbf{P}(X)$, then $g^\Delta \leq g$,
- (2) for any disjunct g in S , $v_1(g) \geq (c_{i-1}, f)$ or $v_1(g) \leq \alpha$,

then the following two statements are equivalent:

- (i) there exists an α -QLS resolution deduction (w.r.t. v_1) from S to $\alpha \ominus$,
- (ii) there exists an c_{n-i+1} -QLS resolution deduction (w.r.t. v_2) from S to $c_{n-i+1} \ominus$,

where v_2 is the valuation in $\mathcal{L}_{V_n}\mathbf{P}(X)$ that satisfies Convention 3.1.

Proof. The proof is similar to that of Theorem 3.2.

Example 3.2. Let $C_1 = x \rightarrow y$, $C_2 = (x \rightarrow z)' \vee (s \rightarrow u)$, $C_3 = y' \vee (y \rightarrow z) \vee (s \rightarrow (Al, f))$, $C_4 = (s \rightarrow u)' \vee (r \rightarrow (So, t))'$, $C_5 = r \rightarrow (Ex, f)$ be generalized clauses in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(9 \times 2)}\mathbf{P}(X)$ and $S = C_1 \wedge C_2 \wedge C_3 \wedge C_4 \wedge C_5$, where $(Al, f), (So, t), (Ex, f) \in L_{V(9 \times 2)}$ and x, y, z, r, s, u are propositional variables. If $\alpha = (Qu, f)$, then there exists an α -QLS resolution deduction from S to $\alpha \ominus$.

In fact, for any disjunct g occurring in S , $g^\Delta \leq g$ holds obviously, i.e., the condition (1) of Theorem 3.5 holds. As there exists a valuation v_1 in $\mathcal{L}_{V(9 \times 2)}\mathbf{P}(X)$ such that $v_1(x) = (Ab, t)$, $v_1(y) = (Hi, f)$, $v_1(z) = (So, t)$, $v_1(s) = (Ex, f)$, $v_1(t) = (Ve, f)$, $v_1(r) = (Al, f)$, so we can obtain that the condition (2) of Theorem 3.5 also holds. Hence, according to Theorem 3.5, we just need to prove that there exists a Qu -QLS resolution deduction (w.r.t. v_2) from S to $Qu \ominus$. Let S have the following lock:

- $C_1 = {}_1(x \rightarrow y)$,
- $C_2 = {}_2(x \rightarrow z)' \vee {}_5(s \rightarrow u)$,
- $C_3 = {}_3 y' \vee {}_4(y \rightarrow z) \vee {}_6(s \rightarrow (Al, f))$,
- $C_4 = {}_7(s \rightarrow u)' \vee {}_8(r \rightarrow (So, t))'$,
- $C_5 = {}_9(r \rightarrow (Ex, f))$.

Since v_2 is the valuation in $\mathcal{L}_{V_9}\mathbf{P}(X)$ such that $v_2(x) = Ab$, $v_2(y) = So$, $v_2(z) = So$, $v_2(s) = Ex$, $v_2(t) = Ra$, $v_2(r) = Qu$, so we have $v_2(C_1) < Qu$, $v_2(C_4) < Qu$, $v_2(C_2) > Qu$, $v_2(C_3) > Qu$, $v_2(C_5) > Qu$. Furthermore, there exists the following Qu -QLS resolution deduction (w.r.t. v_2) ω^* :

- (1) ${}_1(x \rightarrow y)$
- (2) ${}_2(x \rightarrow z)' \vee {}_5(s \rightarrow u)$
- (3) ${}_3 y' \vee {}_4(y \rightarrow z) \vee {}_6(s \rightarrow (Al, f))$
- (4) ${}_7(s \rightarrow u)' \vee {}_8(r \rightarrow (So, t))'$
- (5) ${}_9(r \rightarrow (Ex, f))$
- (6) ${}_8(r \rightarrow (So, t))' \vee Qu$ by (1), (2), (3), (4)
- (7) Qu by (5), (6).

Therefore, there exists a Qu -QLS resolution deduction (w.r.t. v_2) from S to $Qu\text{-}\odot$. Furthermore, we can obtain an α -QLS resolution deduction ω (w.r.t. v_1) from S to $\alpha\text{-}\odot$ by replacing all Qu occurring in ω^* with α .

Combining with Theorems 3.2 and 3.5, we can draw the conclusion that α -QLS resolution for linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ can be equivalently transformed into that for lattice-valued propositional logic $\mathcal{L}_{V_n}P(X)$ under some conditions, which dramatically reduces the difficulty of α -resolution automated reasoning for linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$.

4 Conclusion

α -Quasi-lock semantic resolution method in linguistic truth-valued lattice-valued propositional logic $\mathcal{L}_{V(n \times 2)}P(X)$ is analyzed in the present paper. Concretely, α -quasi-lock semantic resolution in $\mathcal{L}_{V(n \times 2)}P(X)$ is equivalently transformed into that in a more simple lattice-valued propositional logic $\mathcal{L}_{V_n}P(X)$.

Acknowledgments. This work is partially supported by the National Natural Science Foundation of P. R. China (Grant No. 60875034), the Applied Basic Research Projects in Sichuan Province of P. R. China (Grant No. 2011JY0092), the projects TIN-2009-0828, P08-TIC-3548 and FEDER funds.

References

1. Luckham, D.: Refinements in resolution theory. In: Proc. IRIA Symp. Automatic Demonstration, pp. 163–190. Springer, New York (1970)
2. Loveland, D.W.: A linear format for resolution. In: Proc. IRIA Symp. Automatic Demonstration, pp. 147–162. Springer, New York (1970)
3. Robinson, J.P.: A machine-oriented logic based on the resolution principle. J. ACM 12, 23–41 (1965)
4. Slagle, J.R.: Automatic theorem proving with renamable and semantic resolution. J. ACM 14(4), 687–697 (1967)
5. Li, X.B., Qiu, X.P., Chang, Z.Y., Xu, Y.: The properties of α -resolution and J-resolution based on lattice-valued propositional logic system. Chinese Quarterly Journal of Mathematics 6, 262–269 (2008) (in Chinese)
6. Zhong, X.M., Xu, Y., Liu, J., Ruan, D., Chen, S.W.: General form of α -resolution based on linguistic truth-valued lattice-valued logic (finished)
7. Zhong, X.M., Xu, Y., Liu, J., Ruan, D., Chen, S.W.: α -Quasi-lock semantic resolution method based on lattice-valued logic (finished)
8. Zhong, X.M., Xu, Y.: α -Quasi-lock semantic resolution method in lattice-valued logic based on lattice implication algebra $L_{n \times 2}$ (finished)

9. He, X.X., Xu, Y., Li, Y.F., Liu, J., Martinez, L., Ruan, D.: α -Satisfiability and α -Lock Resolution for a Lattice-Valued Logic LP(X). In: The 5th International Conference on Hybrid Artificial Intelligence Systems, San Sebastián, Spain, June 23-25, pp. 320–327 (2010)
10. Xu, Y., Ruan, D., Kerre, E.E., Liu, J.: α -Resolution principle based on lattice-valued propositional logic LP(X). *Inform. Sci.* 130, 195–223 (2000)
11. Xu, Y., Ruan, D., Qin, K.Y., Liu, J.: *Lattice-Valued Logic*. Springer, Heidelberg (2003)
12. Xu, Y., Chen, S.W., Ma, J.: Linguistic truth-valued lattice implication algebra and its properties. In: Proc. IMACS MuLticference on “Computational Engineering in Systems Applications” (CESA 2006), Beijing, China, pp. 1413–1418 (2006)
13. Xu, Y., Zhong, X.M., et al.: General form of a-resolution principle based on lattice-valued logic with truth-value in lattice implication algebras (to be published)

A Real-Time Software GPS Receiver Based on MAX2769

Xiao Xu and Yijin Chen

College of Geoscience and Surveying Engineering, China University of Mining & Technology(Beijing), Beijing 100083, China

xuxiao1028@yahoo.com.cn, y.j.chen@263.net

Abstract. Recently more and more people pay attention to software GPS receiver, because it provides a high level of flexibility and simplicity for many potential applications. The resulting software GPS receivers offer considerable flexibility in modifying settings to accommodate new applications without re-designing hardware, choosing an IF frequency, or implementing future upgrades. This paper introduces the design of a software-based receiver for L1-band civilian GPS applications, including the design of the RF front end based on MAX2769, the algorithms for the signal acquisition, tracking and positioning decoding, which are demonstrated.

Keywords: software GPS receiver, RF front end, acquisition, tracking, pseudorange.

1 Introduction

Global Positioning System (GPS) is a satellite-based navigation system that is widely used in civil and military location. The most important part in this system is the receiver. The modern GPS receivers are based on application specific integrated circuit (ASIC) for signal processing and fast microprocessors for application calculations. But the ASIC can't easily be reprogrammed like a microprocessor. When wanting to change the arithmetic, the ASIC must be redesigned by the manufacturer and this is costly. With the development of GPS technology, for example, adding L5 band signal, and to the special use receiver, which can be compatible with various satellite navigation and positioning system, intermediate frequency digital signal processing algorithm often need to be changed. But the special chips of hardware receiver can't meet these requirements [7].

The software radio is modern radio receiver and transmitter concept. It is widely used in software GPS receiver design for its flexibility and versatility [2]. According to the implemented software, the receiver can serve as a single system receiver with advanced algorithms of signal processing or a combined receiver for various systems. Implemented algorithms can enhance performance of the signal processing and position determination [1].

This paper describes the design of the hardware and software of a real-time software GPS receiver, based on MAX2769 RF front end. The MAX2769 transfers raw recovered data to the PC host, which executes the baseband decoding in software.

2 GPS Signal

A GPS system consists of at least 24 satellites that operate on the L1:1575.42 MHz and L2:1227.6 MHz frequency, a ground-control station, and user equipment (receivers). All current satellites transmit a public Coarse Acquisition code (C/A) on L1 and an encrypted P(Y) code on L1 and L2. Each satellite is assigned a unique ID code that it was used to identify its signal transmission. In order to make a position determination, a receiver must identify the code and then synchronize a local replica of it using at least four satellites.

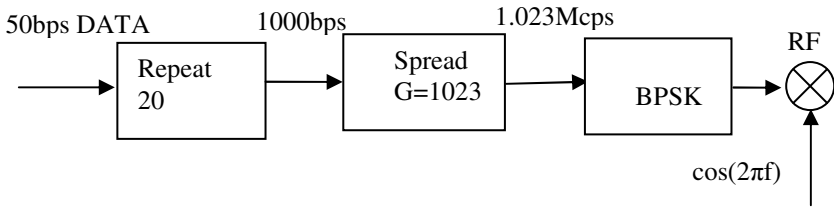


Fig. 1. iA spread-spectrum approach for sending the signal

The navigation data transmitted by the satellites are Binary Phase Shift Keying (BPSK) encoded at 50 bps and broadcast using a direct-sequence spread-spectrum technique known as Code Division Multiple Access (CDMA). Figure 1 shows the signal generation block for civilian applications. First, the 50bps navigation message is repeated 20 times to produce a 1000bps bit stream. The repeated signal is then spread by a unique C/A code with a length of 1023 chips—a chip is the rate at which the pseudorandom-noise (PRN) code is applied. The result is a baseband signal of 1.023Mcps.

3 RF Front End Architecture

Because of the received signal's analogue nature, the software receiver for processing of the real signals always consists of an analogue part called RF front end. The signal in the output of the front end is converted to the digital domain and processed in a programmable digital signal processor, like DSP, FPGA and PC. The performance of the software receiver depends on the performance of the receiver's front end hardware.

The RF front end of a software GPS receiver first amplifies the weak incoming signal with a low-noise amplifier (LNA), and then down converts the signal to a low intermediate frequency (IF). This down conversion is accomplished by mixing the input RF signal with the local oscillator signal using one or two mixers. The resulting analog IF signal is converted to a digital IF signal by the analog-to-digital converter (ADC) [5].

crystal/reference input. If the desired load value is between 11pF and 17pF, the value of this coupling capacitor can be made large, so as not to affect the programmed value. For crystals with load capacitances below 11pF, the coupling capacitor can be made small to add in series with the internal bank, reducing the load seen at the device. Either way, the final frequency trimming can be done internally through the SPI interface.

An excellent noise figure of 1.4dB for the cascaded chain allows this device to be used with a passive antenna, as well as an active antenna can be used in an automotive application. A voltage is provided at pin 3 specifically to bias the active device. This voltage can be turned off through the SPI interface for passive-antenna applications. If, however, the voltage is enabled, then LNA selection can be done automatically depending on whether there is an active antenna present. In the LNA-gated mode, the receiver is configured to automatically switch between the two LNAs contingent on whether a load current in excess of 1.5mA is detected at the antenna bias pin. A user does not need separate designs for applications using active and passive antennas; the chip automatically selects the appropriate LNA for any application. If automatic LNA selection is not desired, it also can be disabled.

The design accommodates a wide range of reference frequencies between 8MHz and 44MHz with a default setting of 16.328MHz. The IF frequency is adjustable in 63 steps between 0 and 12.5MHz, with our used setting of 4.092MHz.

Most GPS devices provide only a single-output mode. The MAX2769's output can be programmed to be analog, CMOS, or limited differential logic in unsigned or complementary binary format with 1- to 3-bit output from the ADC. Here we choose CMOS, with 2-bit output. A 3-wire (SPI) digital bus, including SDATA (pin 8), SCLK (pin 9) and CS/ (pin 10) is used to program the registers of the MAX2769.

The DC power consumption of this circuit is very low. Power required is typically 16mA to 23mA at 3V. Using SPI control, the device can be placed in idle mode, in which only the clock buffer and temperature sensor are active and the current consumption drops to 0.5mA. The overall gain from RF input to IF output can be tuned or automatically controlled over a 60dB to 115dB range.

In order to transmit the data, coming from the front end, to PC, through an industry-standard USB 2.0 interface, we choose an intermediate frequency signal acquisition board designed by Hellognss Company. The software baseband program uses the input to calculate a position fix and, subsequently, to perform tracking.

4 Signal Acquisition

In order to track and demodulate the GPS navigation information, capture must be used to get the visible GPS satellite signals. In this process, GPS receivers must achieve and maintain lock on carrier Doppler shift and the pseudo-random number (PRN) code phase. A standard GPS receiver accomplishes this with two separate, consecutive operations: acquisition and tracking [4]. Acquisition searches for initial estimates of Doppler shift and PRN code phase. The tracking algorithm then uses these estimates to initiate a delay-locked loop (DLL), and a phase-locked loop (PLL), to respectively keep the replicas of the code and carrier signal aligned with those of the received signal.

The purpose of acquisition is to determine visible satellites and coarse values of carrier frequency and code phase of every satellite's signal. The output of acquisition is a collection of tracking objects for each signal transmitter to be tracked. To identify satellites, 32 different PRN sequences are used. The code phase is the time alignment of PRN code in the current block of data. When the local generated PRN code is perfectly aligned with the incoming code, the incoming code can be removed from the signal.

In the receiver, the GPS signal is first down converted to in-phase and quadrature (I and Q) components. A pair of I/Q correlators then correlate the I and Q baseband signals with the locally generated PRN sequence [6]. After integrating over the duration of one bit, the I-Q correlator outputs are summed to provide an output-decision variable. Whenever the decision variable exceeds a certain threshold value, the system assumes that the corresponding acquisition was successful and proceeds to tracking mode. Otherwise, the relative phase of the locally generated PRN sequence and the oscillator frequency are adjusted to update the decision variable, and the above process is repeated [9].

In software GPS receiver we use the method called parallel code phase search acquisition. This method combines the Doppler-frequency and code-phase searches into one search, which, after a fast Fourier transform (FFT) of the PRN code, reflects all code-phase information into the frequency domain. We then need only to search the space over the Doppler-frequency offset, thereby implementing a fast and effective software search [8].

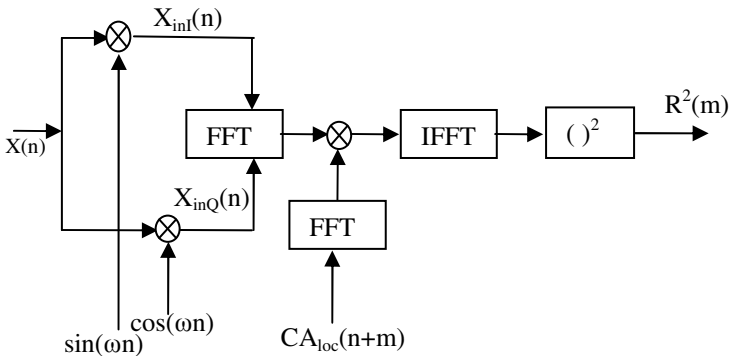


Fig. 3. Parallel Code Phase Search Acquisition

Figure 3 shows the process of parallel code phase search acquisition. The system implements this search by first multiplying the incoming signal with the locally generated sine and cosine carrier waves. The I and Q components are then combined as a complex input to an FFT block. The result of this Fourier transform is multiplied with the conjugate of a PRN code's FFT transform. Then, the product of the incoming signal and local code, which represents the correction between the incoming and carrier frequencies, is applied to an inverse Fourier transform whose squared output feeds back to the decision logic. Finally, observe the results, and compare the maximum value with the predetermined threshold. If the maximum value is greater than the

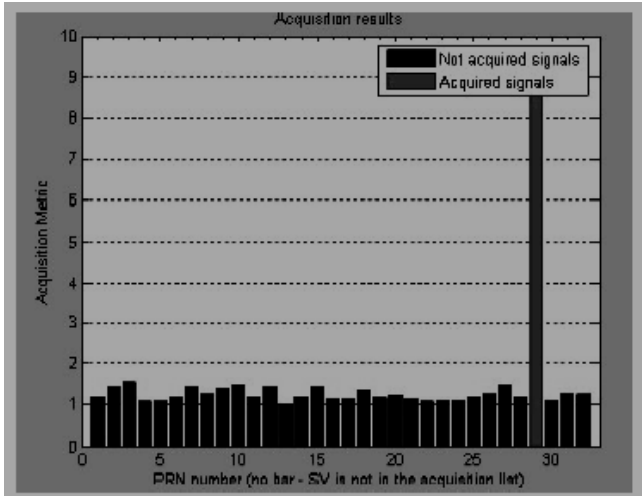


Fig. 4. Capture the signal

threshold already be set, signal acquisition is successful, otherwise change the Doppler frequency and repeat the process until capture the signal [10], as shown in figure 4.

5 Tracking Refines Alignment

After acquisition, the frequency and code offset parameters of a satellite signal are roughly known. The purpose of tracking, therefore, is to refine this alignment so that the system can demodulate the data with exact code-phase and frequency information for the specific satellite. This is done by wiping off the carrier wave from the signal. Next, the signal is multiplied with a code replica and the output of this multiplication gives the navigation message. So the tracking module has to generate two replicas, one for the carrier and one for the code. And correspondingly tracking includes code-phase tracking and carrier-frequency tracking [9].

The goal for a code tracking loop is to keep track of the code phase of a specific code in the signal. The code tracking loop in the GPS receiver is a delay lock loop (DLL) called an Early-Late tracking loop [4], as shown in Figure 5.

To demodulate the navigation data successfully, an exact carrier wave replica has to be generated. To track a carrier wave signal, phase lock loop (PLL) or Costas loop is often used.

A Costas loop, as shown in figure 6, is a phase-locked loop, which is insensitive to phase transition. The carrier frequency loop receives a cw signal phase modulated only by the navigation data as the C/A code is stripped off from the input signal. The acquisition program determines the initial value of the carrier frequency. The voltage-controlled oscillator (VCO) generates a carrier frequency according to the value obtained from the acquisition program. This signal is divided into two paths: a direct one and one with a 90-degree phase shift. These two signals are correlated with the input signal. The outputs of the correlators are filtered and their phases are compared

against each other through an arctangent comparator [11]. The output of the comparator is filtered again and generates a control signal. This control signal is used to tune the oscillator to generate a carrier frequency to follow the input cw signal. This carrier frequency is also used to strip the carrier from the input signal [8].

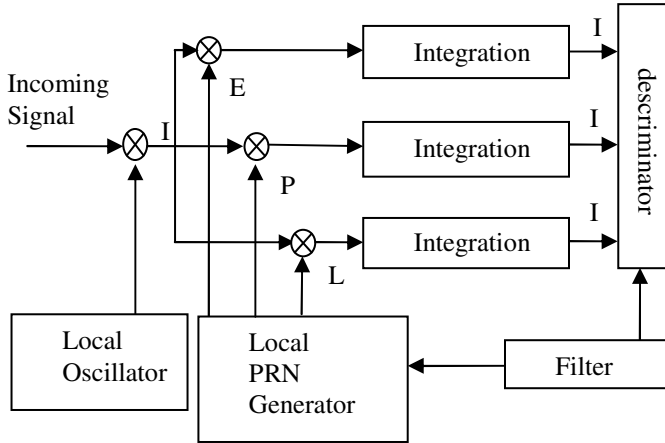


Fig. 5. DDL code tracking

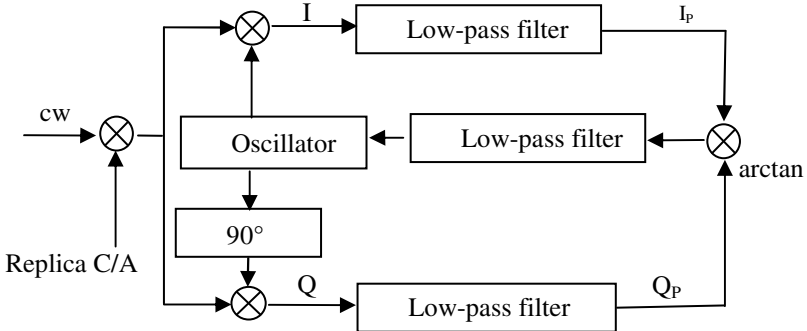


Fig. 6. Costas Loop

As it is shown in figure 7, getting the last one graphics, means that the collected data contains satellite signal and the data is continuous.

After acquisition and tracking the initial synchronization has to be established. The value of the in-phase arm of tracking block truncated to the values of 1 and -1[9], which is the result of the output from the tracking loop. The tracking block supplies this output every millisecond. The signal from the tracking block must be converted from every 1000 bps to 50 bps. For bit synchronization, we first need to identify the

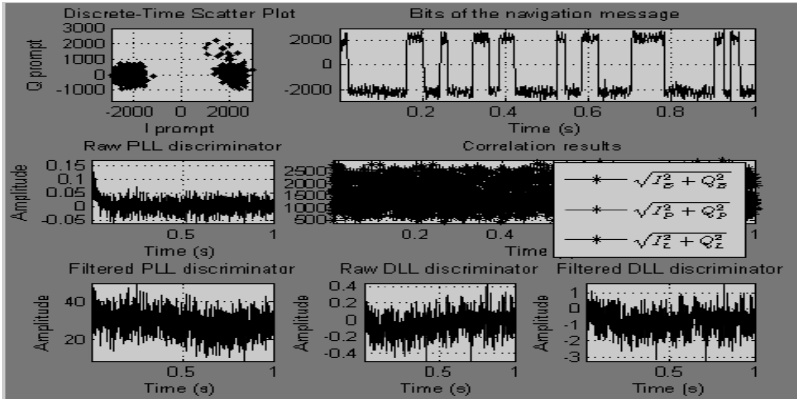


Fig. 7. Tracking result

beginning of a bit in time. This is accomplished by finding the zero-crossing edge, which indicates the beginning of a bit. When that edge is known, we can partition the 1000bps input stream at 20ms intervals, knowing that the duration of a navigation data message is 20ms. Finally, the bit samples in a 20ms interval are summed and averaged to decode the navigation data [6].

6 Positioning Decoding

Once the tracking is achieved, the output will be converted into navigation data through subframe matching and parity checking. From the subframes the ephemeris data such as the week number can be found. The position of the satellite can be determined from the ephemeris data. The pseudoranges between the receiver and the satellites can also be determined. Once all the necessary information is obtained, satellite positions and the user position can be calculated. Finally, the user position is put in the desired coordinate system.

In the process above, the most important part is the calculation of pseudorange. Because the local receiver's time can't be got in software GPS receiver, so it is impossible to obtain the transmission time between the satellite and the receiver, and only through relative measurement to obtain pseudorange. All the beginning points of subframe 1 from different satellites are transmitted at the same time except for the clock correction terms of each satellite. As a result one can consider that the subframes from different satellites are transmitted at the same time. Since the beginnings of subframe 1 from different satellites are received at different times, this difference time represents the time (or distance) difference from the satellite to the receiver. Therefore, it represents the relative pseudorange. The relative pseudorange as observed data can be used for positioning decoding [7].

7 Conclusion

Designed on Maxim's advanced, low-power process technology, the MAX2769 offers the highest performance and integration at a low cost. Incorporated on the chip is the complete receiver chain, including a dual-input LNA and mixer, followed by the image-rejected filter, PGA, VCO, fractional-N frequency synthesizer, crystal oscillator, and a ADC, thus significantly reducing the development time for applications. The total cascaded noise figure of this receiver is as low as 1.4dB.

A MAX2769 based RF front end which is suitable for the software GPS receiver and its algorithms have been demonstrated. The integration of portability, high precision and real-time performance has been achieved in the design. The design is part of the GNSS software receivers in relation to the multiple satellites system and multiple frequencies. It provides technical guidance for the future work in GNSS software receivers.

Acknowledgments. The author would like to thank Professor Yijin CHEN for his constant encouragement and guidance.

References

1. Brown, A.: Benefits of Software GPS Receiver for Enhanced Signal Processing. GPS Solutions (2000)
2. Mac Gougan, G., Normark, P.L., Stalberg, C.: Satellite navigation evolution: the software GNSS receiver. GPS World (2005)
3. MAX2769 DATASHEET, <http://www.maxim-ic.com/>
4. Sonowal, N., Yadav, R., Kannan, S.: Real Time GPS Software Receiver with New Fast Signal Tracking Method. In: Proceedings of 2008 IEEE Radio and Wireless Symposium, Orlando, FL, USA (2008)
5. Weber, D., Bremer, R.: Universal GPS Receiver Lets You Use a Laptop PC for Soft Baseband Processing. Maxim's Engineering Journal 63 (2008)
6. Jin, T., Qin, H., Zhu, J.: Real-time GPS Software Receiver Correlator Design. Communications and Networking in China (2007)
7. Tsui, J.B.-Y.: Fundamentals of Global Positioning System Receivers: A Software Approach, 2nd edn. John Wiley & Sons, Inc., New York (2004)
8. Li, Z., Hao, J., Li, J., Zhang, C.: IF Signal Processing Algorithm for GPS Software Receiver. Congress on Image and Signal Processing 1 (2008)
9. Molino, A., Nicola, M., Pini, M., Fantino, M.: N-Gen GNSS Software Receiver for Acquisition and Tracking Algorithms Validation. In: 17th European Signal Processing Conference (2009)
10. Borio, Driscoll, D. O., Lachapelle, C.: Composite GNSS Signal Acquisition over Multiple Code Periods. Aerospace and Electronic Systems (2010)
11. Huang, Y., Tang, X.: Differential Correlator for Tracking GPS Signals. Information Engineering and Computer Science (2009)

Duality in Lattice Implication Algebra

Li Zhao and Yang Xu

College of Mathematics, Southwest Jiaotong University,
610031 Chengdu, China
yuyehuaixiang@163.com,
xuyang@home.swjtu.edu.cn

Abstract. According to the general form of principle of duality in the sense of class [1], this paper tries to study the dual operators of operators in lattice implication algebra [2], especially the dual operator of implication operator and gives the expression of principle of duality in lattice implication algebra. Then we will show its value in developing theories in corresponding field.

Keywords: Lattice implication algebra, Principle of duality, Filter, LI-ideal.

1 Introduction

In order to provide a logic basic for uncertainty reasoning and automated reasoning, Xu [2] proposed lattice implication algebra depending on lattice and implication algebra in 1993. After that, many scholars have been researching it actively, boosting its development greatly. In 1993, Xu and Qin [3] proposed the concept of filter and studied its properties. In 1995, Xu and Qin [4] proposed the concepts of fuzzy filter and fuzzy implicative filter and studied their properties. In 1996, Liu and Xu [5] proposed the concepts of I-feature filter (I-filter) and involutory filter (involution filter) and studied their properties. In 1998, Meng [6] proposed the concept of prime filter and studied its properties. In 1998, Y.B. Jun etc. [7] proposed the concept of LI-ideal and studied its properties. In 1999, Y.B. Jun and Xu [8] proposed the concept of fuzzy LI-ideal and studied its properties. In 1999, Y.B. Jun [9] proposed the concept of prime LI-ideal and studied its properties. In 2003, Y. L. Liu etc. [10] proposed the concepts of implicative LI-ideal and maximal LI-ideal and studied their properties. In 2009, Lai [11] proposed the concepts of weak filter and weak LI-ideal and studied their properties and so on.

Since the filter and ideal of lattice are dual [12], the duality of filter and LI-ideal of lattice implication algebra may probably exist. What's more, every concrete kind of filters of lattice implication algebra may be dual with corresponding LI-ideals as well. Bai [13] studied the embodiments of duality in lattice implication algebra in terms of pairs of theorems and gave some dual theorems of the known ones. But the details of the dual process haven't been studied. If we master the dual method in lattice implication algebra, we can easily get a new concept or theorem from the known one. And the new one may be convenient for researching. This paper will try to show the details of the dual process by giving the principle of duality in lattice implication algebra.

The structure of this paper is organized as follows: in section 2, the preliminaries are shown; in section 3, the dual operators of those in lattice implication algebra are discussed and the principle of duality in lattice implication algebra is given; in section 4, the applied value of principle of duality in lattice implication algebra in enriching concepts and theorems is shown by several examples; finally, the conclusions are given.

2 Preliminaries

Definition 2.1 [14] (Lattice implication algebra). A bounded lattice (L, \vee, \wedge, O, I) with an ordered-reversing involution $'$, O and I the smallest and the greatest element of L respectively, and a binary operator $\rightarrow: L \times L \rightarrow L$ is called a lattice implication algebra if for any $x, y, z \in L$, it satisfies the following conditions:

$$\begin{aligned} (I_1) & x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z); \\ (I_2) & x \rightarrow x = I; \\ (I_3) & x \rightarrow y = y' \rightarrow x'; \\ (I_4) & x \rightarrow y = y \rightarrow x = I \text{ implies } x = y; \\ (I_5) & (x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x; \\ (L_1) & (x \vee y) \rightarrow z = (x \rightarrow z) \wedge (y \rightarrow z); \\ (L_2) & (x \wedge y) \rightarrow z = (x \rightarrow z) \vee (y \rightarrow z). \end{aligned}$$

Remark. Lattice implication algebra is always written as LIA in the following of this paper.

Principle 2.1 [1] (General form of principle of duality in the sense of class). *Assume that L is a class, in which the objects have the same nature (for example, they are all lattices, partial sets or categories and so on) and are all systems (that means every object in the class not only contains elements but also operators). There is a convolution transform D in L (that means D^2 is an identical transformation), and D has certain finite-transformation rules or procedure (that's to say it can be realized by using D finite times to transform the elements and operators of an object in the class). L is closed under D , which means "for any $A \in L$, we can get $D(A) \in L$ ", here $D(A)$ (denoted by A°) is called the duality of A .*

Let S be a declarative sentence on L . If the elements and operators of the objects in L which are referred in S , are substituted by their duality, we will get the duality $D(S)$ (denoted by S°) of S . If the substance described by S is a concept, then S° is its dual concept; if the substance described by S is a proposition, then S° is its dual proposition. Specially, if S is a truth-valued theorem on A and A° , so is S° . Here D should satisfy that " S is a theorem on A " can imply " S° is a theorem on A° ".

The above principle contains two important points. First, it shows how to get the duality of a concept or a theorem. Second, it shows that the theorem S° obtained by dual transformation from S has the same truth-value with S .

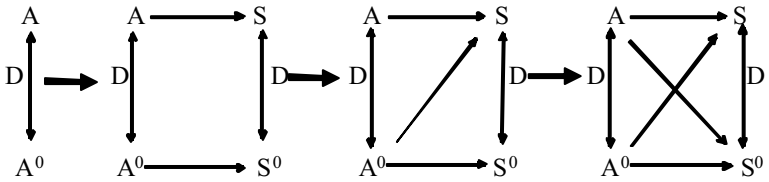


Fig. 1. Schematic diagram of the “general form of principle of duality in the sense of class”[1]

In Figure 1, A and A° are two dual objects in a class L ; “ $A \rightarrow S$ ” means S is a declarative sentence on A ; “ $A^\circ \rightarrow S^\circ$ ” means S° is a declarative sentence on A° . This figure shows that the duality of objects A and A° leads to the duality of their properties and if the property of A is also that of A° , then the property of A° is also that of A .

Axiom 2.1 [15] (Algebraic substitution axiom). *In any algebraic identical equation, every letter symbol just is a general variable. So, it can be substituted by the other letters or identical function expression (just it is significant after substitution). After substitution, the identical equation still holds.*

Remark. Other relevant information can be seen in corresponding references, for example, the definitions of implication homomorphism, $\ker(f)$, $D\text{-ker}(f)$, filter generated by a set and LI-ideal generated by a set can be consulted in reference [14].

3 Principle of Duality in LIA

3.1 Dual Operators in LIA

Definition 3.1.1 (Duality of implication operator). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA and $\mapsto: L \times L \rightarrow L$ be a mapping, \mapsto is called the dual operator of \rightarrow if for any $x, y \in L, x \mapsto y = (y \rightarrow x)'$.

\mapsto defined above is indeed the dual operator of \rightarrow . The following is the reason: let D be the convolution transform (substitute the elements and operators by their duality), then

$$D(D(x \rightarrow y)) = D(x' \mapsto y') = D((y' \rightarrow x')') = (y \mapsto x)' = ((x \rightarrow y)')' = x \rightarrow y.$$

This implies that \mapsto is the dual operator of \rightarrow .

For the same reason, “ \vee ” and “ \wedge ”, “ \geq ” and “ \leq ”, “ \otimes [14]” and “ \oplus [14]” are all dual operators. Moreover, “ $'$ ” is self-dual operator.

Definite 3.1.2 (Duality of membership function and non-membership function). Let $A = \{ (x, \alpha_A(x), \beta_A(x)) \mid x \in X \}$ be an intuitionistic fuzzy set on a non-empty set X , in which $\alpha_A : X \rightarrow [0,1]$ is the membership function, $\beta_A : X \rightarrow [0,1]$ is the non-membership function and $0 \leq \alpha_A(x) + \beta_A(x) \leq 1$. Then $\alpha_A'(x), \beta_A'(x)$ are said to be the dual functions of $\alpha_A(x)$ and $\beta_A(x)$, if $\alpha_A'(x) = 1 - \alpha_A(x), \beta_A'(x) = 1 - \beta_A(x)$.

Because the concept “intuitionistic fuzzy set” is the extension of “fuzzy set”, the duality of membership function in fuzzy set follows the same rule.

3.2 Principle of Duality in LIA

Principle 3.2.1 (Principle of duality in LIA). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA, P be a concept of (prime, implicative, fuzzy...) filter and for any $a_i \in L, i \in N^+, P = (a_1, a_2, \dots, \vee, \wedge, ', \rightarrow, \geq, \leq, \min, \max, \alpha_A(x), \beta_A(x))$. If $P' = (a_1', a_2', \dots, \wedge, \vee, ', \rightarrow, \leq, \geq, \max, \min, \alpha_A'(x), \beta_A'(x))$, P' is said to be the concept of (prime, implicative, fuzzy...)LI-ideal. That's to say P' is the dual concept of P .

Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA, A and A' be two dual concepts in \mathcal{L} , P be a theorem on A or A' , and for any $a_i \in L, i \in N^+, P = (a_1, a_2, \dots, A, A', \vee, \wedge, ', \rightarrow, \geq, \leq, \min, \max, \alpha_A(x), \beta_A(x))$ or $P = (a_1, a_2, \dots, A, \vee, \wedge, ', \rightarrow, \geq, \leq, \min, \max, \alpha_A(x), \beta_A(x))$. Accordingly if $P' = (a_1', a_2', \dots, A', A, \wedge, \vee, ', \rightarrow, \leq, \geq, \max, \min, \alpha_A'(x), \beta_A'(x))$ or $P' = (a_1', a_2', \dots, A', \wedge, \vee, ', \rightarrow, \leq, \geq, \max, \min, \alpha_A'(x), \beta_A'(x))$, P' is said to be the dual theorem of P .

4 Application of Principle of Duality in LIA

4.1 Enriching Concepts

Applying the first part of principle of duality in LIA, we can easily get the concepts in the second column of Table 1 from those in the first. During the course of transformation, “algebraic substitution axiom” is usually necessary.

Table 1. List of some dual concepts in LIA

Original Concept	Dual Concept
Lattice filter	Lattice ideal
Filter	LI-ideal
Implicative filter	Implicative LI-ideal
Prime filter	Prime LI-ideal
Weak filter	Weak LI-ideal
Fuzzy filte	Fuzzy LI-ideal
Fuzzy implicative filter [14]	Fuzzy implicative LI-ideal [16]
Intuitionistic fuzzy filter [17]	Intuitionistic fuzzy LI-ideal [14]

Example 1. Getting the concept of implicative LI-ideal from that of implicative filter.

Definition 4.1.1 [5] (Implicative filter). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA and $J \subseteq L$. J is called an implicative filter of \mathcal{L} if for any $x, y, z \in L$, the following two conditions can be satisfied:

$$(1) I \in J; (2) x \rightarrow (y \rightarrow z) \in J, x \rightarrow y \in J \Rightarrow x \rightarrow z \in J.$$

Definition 4.1.2 [10] (Implicative LI-ideal). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA and $A \subseteq L$. A is called an implicative LI-ideal of \mathcal{L} if for any $x, y, z \in L$, the following two conditions can be satisfied:

$$(1) O \in A; (2) (((x \rightarrow y)' \rightarrow y)' \rightarrow z)' \in A, z \in A \Rightarrow (x \rightarrow y)' \in A.$$

Through the following dual process, we can easily get the concept of implicative LI-ideal in Definition 4.1.2'. Here, we should note that after applying the “principle of duality in LIA” and “algebraic substitution axiom” to Definition 4.1.1, “ x, y, z ” in the precondition remain unchanged.

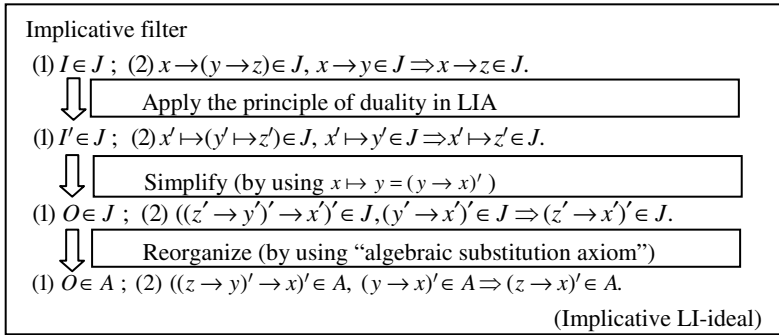


Fig. 2. Schematic diagram of transformation from Definition 4.1.1 to 4.1.2'

Definition 4.1.2' (Implicative LI-ideal). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA and $A \subseteq L$. A is called an implicative LI-ideal of \mathcal{L} if for any $x, y, z \in L$, the following two conditions can be satisfied:

$$(1) O \in A; (2) ((z \rightarrow y)' \rightarrow x)' \in A, (y \rightarrow x)' \in A \Rightarrow (z \rightarrow x)' \in A.$$

Theorem 4.1.1. As the definition of implicative LI-ideal of LIA, Definition 4.1.2' is equivalent to Definition 4.1.2.

Proof. Obviously, the first conditions of the two definitions are equivalent. Let's pay attention to the second conditions of the two definitions.

If Definition 4.1.2 is right, we assume

$$((z \rightarrow y)' \rightarrow x)' \in A, (y \rightarrow x)' \in A.$$

Since

$$(((x \rightarrow y)' \rightarrow y)' \rightarrow z)' \in A, z \in A \Rightarrow (x \rightarrow y)' \in A,$$

let $y = O$, then

$$(x \rightarrow z)' \in A, z \in A \Rightarrow x \in A.$$

That's to say, A is a LI-ideal of \mathcal{L} .

It's obvious that the following formula holds.

$$\begin{aligned} & (((z \rightarrow x)' \rightarrow x)' \rightarrow (y \rightarrow x)')' \rightarrow ((z \rightarrow y)' \rightarrow x)')' \\ &= (((y \rightarrow x) \rightarrow ((z \rightarrow x)' \rightarrow x))' \rightarrow ((z \rightarrow y)' \rightarrow x)')' \\ &= (((z \rightarrow y)' \rightarrow x) \rightarrow ((y \rightarrow x) \rightarrow ((z \rightarrow x)' \rightarrow x)))' \\ &= ((y \rightarrow x) \rightarrow (((z \rightarrow y)' \rightarrow x) \rightarrow ((z \rightarrow x)' \rightarrow x)))' \\ &= ((y \rightarrow x) \rightarrow ((x' \rightarrow (z \rightarrow x)) \vee (z \rightarrow x) \vee (y \rightarrow x)))' \\ &= I' = O \in A \end{aligned}$$

Since

$$((z \rightarrow y)' \rightarrow x)' \in A$$

and A is a LI-ideal of \mathcal{L} , we get

$$(((z \rightarrow x)' \rightarrow x)' \rightarrow (y \rightarrow x)')' \in A.$$

Since $(y \rightarrow x)' \in A$, A is defined as Definition 4.1.2, we have

$$(z \rightarrow x)' \in A.$$

Hence

$$((z \rightarrow y)' \rightarrow x)' \in A, (y \rightarrow x)' \in A \Rightarrow (z \rightarrow x)' \in A,$$

namely Definition 4.1.2' is right.

Conversely, if the Definition 4.1.2' is right, we assume that

$$(((x \rightarrow y)' \rightarrow y)' \rightarrow z)' \in A, z \in A.$$

Since

$$((z \rightarrow y)' \rightarrow x)' \in A, (y \rightarrow x)' \in A \Rightarrow (z \rightarrow x)' \in A,$$

let $x = O$, then

$$(z \rightarrow y)' \in A, y \in A \Rightarrow z \in A.$$

That's to say, A is a LI-ideal of \mathcal{L} , so

$$((x \rightarrow y)' \rightarrow y)' \in A.$$

Since A is defined as Definition 4.1.2' and $(y \rightarrow y)' = O \in A$, we have

$$(x \rightarrow y)' \in A.$$

Therefore

$$(((x \rightarrow y)' \rightarrow y)' \rightarrow z)' \in A, z \in A \Rightarrow (x \rightarrow y)' \in A,$$

namely Definition 4.1.2 is true.

Example 2. Getting the concept of weak LI-ideal from that of weak filter

Definition 4.1.3 [11] (Weak filter). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA and $A \subseteq L, A \neq \emptyset$. Binary operation \oplus is defined as $x \oplus y = x' \rightarrow y$. A is called a weak filter of \mathcal{L} if for any $x, y \in L$, the following condition can be satisfied:

$$x' \oplus y \in A \Rightarrow (x' \oplus y) \oplus x' \in A.$$

Definition 4.1.4 [11] (Weak LI-ideal). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA and $A \subseteq L, A \neq \emptyset$. Binary operation \otimes is defined as $x \otimes y = (x \rightarrow y)'$. A is called a weak LI-ideal of \mathcal{L} if for any $x, y \in L$, the following condition can be satisfied:

$$(x \rightarrow y)' \in A \Rightarrow (x \rightarrow y)' \otimes y' \in A.$$

Applying “principle of duality in LIA” to Definition 4.1.3, which means substituting x, y, \oplus by their dualities x', y', \otimes respectively, we can easily get the concept of weak LI-ideal in Definition 4.1.4'.

Definition 4.1.4' (Weak LI-ideal). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA and $A \subseteq L, A \neq \emptyset$. Binary operation \otimes is defined as $x \otimes y = (x \rightarrow y)'$. A is called a weak LI-ideal of \mathcal{L} if for any $x, y \in L$, the following condition can be satisfied:

$$x \otimes y' \in A \Rightarrow (x \otimes y') \otimes x \in A.$$

Theorem 4.1.2. *As the definition of weak LI-ideal of LIA, Definition 4.1.4' is equivalent to Definition 4.1.4.*

Proof. It is easy to be proved, so omitted.

Example 3. Getting the concept of fuzzy implicative LI-ideal from that of fuzzy implicative filter

Remark. Denote $\mathcal{F}(L)$ as the set of all fuzzy subsets of LIA \mathcal{L} in the following.

Definition 4.1.5 [4] (Fuzzy implicative filter). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA, $A \in \mathcal{F}(L)$ and $A \neq \emptyset$. A is called a fuzzy implicative filter of \mathcal{L} if for any $x, y, z \in L$,

- (1) $A(x) \leq A(I)$; (2) $A(x \rightarrow z) \geq \min\{A(x \rightarrow y), A(x \rightarrow (y \rightarrow z))\}$.

Definition 4.1.6 [16] (Fuzzy implicative LI-ideal). Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA, $A \in \mathcal{F}(L)$ and $A \neq \emptyset$. A is called a fuzzy implicative LI-ideal of \mathcal{L} if for any $x, y, z \in L$,

$$(1) A(x) \leq A(O) ; (2) A((x \rightarrow z)') \geq \min\{A((y \rightarrow z)'), A(((x \rightarrow y)' \rightarrow z)')\}.$$

Through the dual process shown in Figure 3, Definition 4.1.6 can be easily got from Definition 4.1.5.

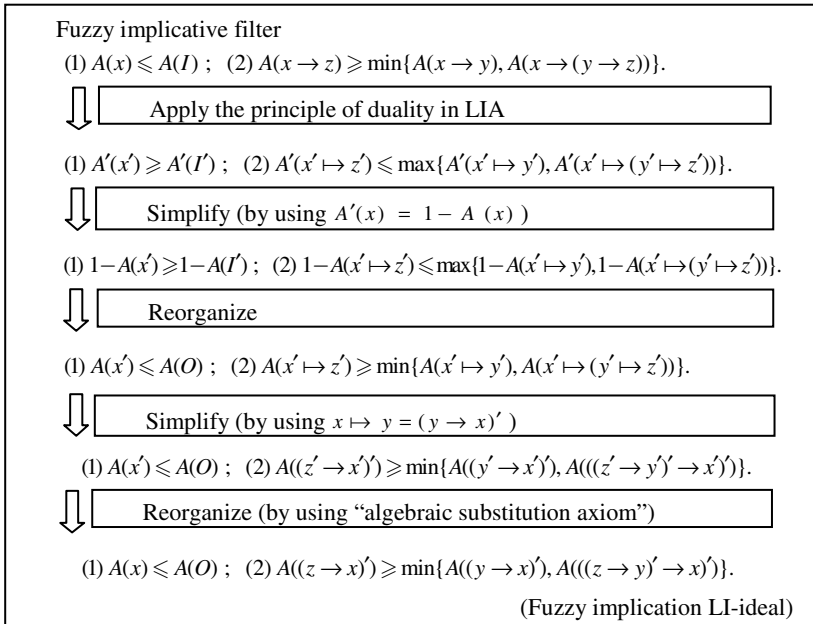


Fig. 3. Schematic diagram of transformation from Definition 4.1.5 to 4.1.6

4.2 Enriching Theorems

Applying the second part of principle of duality in LIA, we can easily get the dualities of certain theorems. The following are two examples.

Example 4

Theorem 4.2.1 [14]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA, $\emptyset \neq A \subseteq L$ and $[A]$ be the filter generated by A . Then

$$[A] = \{ x \mid x \in L, \text{ there exist } a_1, \dots, a_n \in A, n \in \mathbb{N}^+, s. t., [a_1, \dots, a_n, x] = I \}.$$

Theorem 4.2.2 [14]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a LIA, $\emptyset \neq A \subseteq L$ and $\langle A \rangle$ be the LI-ideal generated by A . Then

$$\langle A \rangle = \{ x \mid x \in L, \text{ there exist } a_1, \dots, a_n \in A, n \in N^+, s. t., [a_1', \dots, a_n', x'] = I \}.$$

Obviously, $[A]$ and $\langle A \rangle$ is a pair of dual concepts. Note that

$$[a_1, \dots, a_n, x] = a_1 \rightarrow (a_2 \rightarrow (\dots \rightarrow (a_n \rightarrow x) \dots)).$$

According to the principle of duality in LIA shown in Figure 4, Theorem 4.2.2 which is the duality of Theorem 4.2.1 can be easily got.

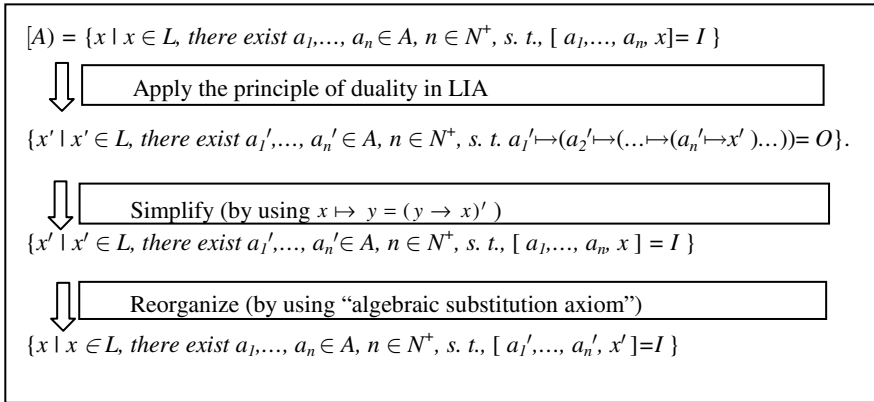


Fig. 4. Schematic diagram of transformation from Theorem 4.2.1 to 4.2.2

Example 5

Theorem 4.2.3 [14]. *Let $\mathcal{L}_1 = (L_1, \vee, \wedge, ', \rightarrow, O_1, I_1)$ and $\mathcal{L}_2 = (L_2, \vee, \wedge, ', \rightarrow, O_2, I_2)$ be LIAs respectively and $f: L_1 \rightarrow L_2$ be an implication homomorphism. If $D\text{-ker}(f) \neq \emptyset$, then $D\text{-ker}(f)$ is a filter of \mathcal{L}_1 .*

Theorem 4.2.4 [14]. *Let $\mathcal{L}_1 = (L_1, \vee, \wedge, ', \rightarrow, O_1, I_1)$ and $\mathcal{L}_2 = (L_2, \vee, \wedge, ', \rightarrow, O_2, I_2)$ be LIAs respectively and $f: L_1 \rightarrow L_2$ be an implication homomorphism. If $\text{ker}(f) \neq \emptyset$, then $\text{ker}(f)$ is a LI-ideal of \mathcal{L}_1 .*

It’s easy to check that $\text{ker}(f)$ and $D\text{-ker}(f)$ is a pair of dual concepts. According to the principle of duality in LIA, if we change “ $D\text{-ker}(f)$ ” and “filter” in Theorem 4.2.3 into “ $\text{ker}(f)$ ” and “LI-ideal”, Theorem 4.2.4 can be got.

5 Conclusion

Through the upper exploration, the duality in LIA has been more distinct. According to the principle of duality in LIA, it is more convenient for us to get a dual concept or theorem from the known one and the new one may be more beneficial to develop the

theory of LIA. For example, as the definition of implicative LI-ideal, Definition 4.1.2' is more concise than Definition 4.1.2. On the other hand, it offers a theory basis for simplifying the study on one of two certain dual concepts.

Acknowledgement. This work is partially supported by the National Natural Science Foundation of P. R. China (Grant No. 60875034) and the Applied Basic Research Projects in Sichuan Province of P. R. China (Grant No. 2011JY0092).

References

1. Xu, B.S., Yin, D.S.: Principle of Duality. *Journal of Qufu Normal University* 16, 96–107 (1990) (in Chinese)
2. Xu, Y.: Lattice Implication Algebras. *Journal of Southwest Jiaotong University* 28, 20–27 (1993) (in Chinese)
3. Xu, Y., Qin, K.Y.: On filters of lattice implication algebras. *The Journal of Fuzzy Mathematics* 1, 251–260 (1993)
4. Xu, Y., Qin, K.Y.: Fuzzy Lattice Implication Algebras. *Journal of Southwest Jiaotong University* 30, 121–127 (1995) (in Chinese)
5. Liu, J., Xu, Y.: On Certain Filters in Lattice Implication Algebras. *Chinese Quarterly Journal of Mathematics* 11, 106–200 (1996)
6. Meng, B.L.: Prime Filters of Lattice Implication Algebras. *Journal of Northwest University (Natural Science Edition)* 28, 189–192 (1998) (in Chinese)
7. Jun, Y.B., Roh, E.H., Xu, Y.: LI-Ideals in lattice implication algebras. *Bull. Korean Mathematics Soc.* 35, 13–24 (1998)
8. Jun, Y.B., Xu, Y.: Fuzzy LI-Ideals in lattice implication algebras. *The Journal of Fuzzy Mathematics* 7, 997–1003 (1999)
9. Jun, Y.B.: On LI-ideals and prime LI-ideals of lattice implication algebras. *Korean Math. Soc.* 36, 369–380 (1999)
10. Liu, Y.L., Liu, S.Y., Xu, Y.: ILI-ideals and Prime LI-ideals in Lattice Implication Algebras. *Information Sciences* 155, 157–175 (2003)
11. Lai, J.J., Xu, Y.: On WLI-ideal Space and the Properties of WLI-ideals in Lattice Implication Algebra. *Appl. Math. Comput.* 31, 113–127 (2009)
12. Birkhoff, G.: *Lattice Theory*, 3rd edn. American Mathematical Society, Providence (1967)
13. Bai, L.J.: Study on the Relation between Lattice Implication Algebras and Correlative Logic Algebras. *Southwest Jiaotong University Master Degree Thesis* (2007) (in Chinese)
14. Xu, Y., Ruan, D., Qin, K.Y., Liu, J.: *Lattice-valued Logic, Lattice-Valued Logic: An Alternative Approach to Treat Fuzziness and Incomparability*. Springer, Berlin (2003)
15. Shi, T.Z.: Algebraic Substitution Axiom and Principle of Duality. *Journal of Chongqing University of Arts and Sciences (Natural Science Edition)* 26, 7–11 (2007) (in Chinese)
16. Zhao, G.F., Xu, Y., Song, Z.M.: Implicative and Fuzzy Implicative Ideals of Lattice Implication Algebras. *Chinese Quarterly Journal of Mathematic* 16, 108 (2001)
17. Pei, Z.: Intuitionistic Fuzzy Filter in Lattice Implication Algebras. *Journal of Xihua University (Natural Science Edition)* 26, 17–20 (2007) (in Chinese)
18. Zhu, Y.Q.: On Implication Algebras Based on Lattices and Their Dual Algebras. *Journal of Mathematical Research and Exposition* 24, 549–555 (2004) (in Chinese)

Kernel Construction via Generalized Eigenvector Decomposition

Yong Liu and Shizhong Liao

School of Computer Science and Technology, Tianjin University, Tianjin, China
szliao@tju.edu.cn

Abstract. Kernel construction is one of the key issues both in current research and application of kernel methods. In this paper, we present an effective kernel construction method, in which we reduce the construction of kernel function to the solutions of generalized eigenvalue problems. Specifically, we first obtain a primal kernel function based on the similarity of instances, and refine it with the conformal transformation. Then we determine the parameters of kernel function by simply solving the generalized eigenvalue problems according to the kernel alignment and Fisher criteria. Our method can avoid local maxima and insure positive semidefinite of the constructed kernel. Experimental results show that our kernel construction method is effective and robust.

Keywords: Kernel Methods, Kernel Construction, Generalized Eigenvalue Problem.

1 Introduction

Kernel methods, such as support vector machine (SVM) [13], kernel principal component analysis (KPCA) [11] and kernel Fisher discriminant analysis (KFDA) [9], have been widely used in pattern recognition and machine learning. The kernel methods map the input space \mathcal{X} into a high dimensional space \mathcal{F} , referred to as kernel feature space, and then build linear machines in the kernel feature space to implement their nonlinear counterparts in the input space. This procedure is usually known as kernelization, in which the inner product of each pair of the mapped data in the kernel feature space is calculated by a kernel function, rather than explicitly using the nonlinear map.

Because of the central role of the kernel, a poor kernel choice can lead to significantly impaired performance. Typically, the practitioner has to select the kernel before learning starts, with common choices being the polynomial kernel and radial basis function (RBF) kernel. The associated kernel parameters can then be determined by optimizing a quality function of kernel [10]. Examples of quality functions include kernel target alignment [6], Fisher [14], generalization error bounds [3], Bayesian probabilities [12] and cross-validation error.

In recent years, there has been a lot of interest on the research of kernel construction [4,7,8]. We note however, the kernels constructed via existing kernel construction methods may not be positive semidefinite and the gradient descent

algorithm adopted by these methods may lead to local maxima. To address the above two problems, we present a new method for constructing kernel function. Specifically, we first decide the primal form of kernel function based on the similarity of the instances. Then, we obtain the final form using the so-called ‘‘conformal transformation’’ on the primal kernel. Finally, based on kernel alignment and Fisher criteria, we obtain the kernel coefficients by simply solving a generalized eigenvalue problem. Compared with the gradient descent algorithm, solving the generalized eigenvalue problem has a number of advantages,

1. It avoids using an iterative algorithm to update alpha (coefficients of the optimal kernel) in each step, since now only a single generalized eigenvalue problem needs to be solved.
2. It avoids running into numerical problems when the Hessian matrices are not well conditioned, and also avoids getting stuck in local maxima.
3. It also avoids tuning parameters such as learning rate and the number of iterations.

The rest of this paper is organized as follows. Section 2 gives the notations and preliminaries. Section 3 describes the new kernel construction method. Experimental results on benchmark sets are presented in Section 4, and the last section is conclusion.

2 Notations and Preliminaries

Let $\mathcal{X} \subset \mathbb{R}^d$ be an input space, $\mathcal{S} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ be labelled instances and $\tilde{\mathcal{S}} = \{\tilde{\mathbf{x}}_i\}_{i=1}^m$ be unlabelled instances, where $\mathbf{x}_i, \tilde{\mathbf{x}}_i \in \mathcal{X}$ and $y_i \in \{-1, +1\}$.

Proposition 1 (Cristianini and Shawe-Taylor [5]). *For any finite set of points $\{\mathbf{x}_i\}_{i=1}^n$, $\mathbf{x}_i \in \mathcal{X}$. Let $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ be a symmetry function on \mathcal{X} . Then K is a kernel function if and only if the matrix*

$$\mathbf{K} = (K(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n \quad (1)$$

is positive semi-definite (has non-negative eigenvalues).

This proposition will be used later to verify whether the function constructed by our method is a kernel function or not. If K is a kernel function, according to [2], we know that there exists the feature mapping $\varphi : \mathcal{X} \rightarrow \mathcal{F}$ such that $K(\cdot, \cdot) = \langle \varphi(\cdot), \varphi(\cdot) \rangle$, where \mathcal{F} is the feature space.

Alignment Criterion. To assess the quality of kernel function, a particular quality function called alignment [6] will be used.

Definition 1 (Alignment). The alignment of kernel K with respect to instances $\mathcal{S} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ is the quantity

$$A(\mathcal{S}, K) = \frac{\langle \mathbf{K}, \mathbf{y}\mathbf{y}^T \rangle_F}{\sqrt{\langle \mathbf{K}, \mathbf{K} \rangle_F \langle \mathbf{y}\mathbf{y}^T, \mathbf{y}\mathbf{y}^T \rangle_F}}, \quad (2)$$

where $\mathbf{y} = (y_1, \dots, y_n)^T$ and $\langle \mathbf{K}, \mathbf{y}\mathbf{y}^T \rangle_F = \sum_{i,j=1}^n K(\mathbf{x}_i, \mathbf{x}_j) y_i y_j$.

This quantity can be viewed as the cosine of the angle between two n^2 dimension vector \mathbf{K} and $\mathbf{y}\mathbf{y}^T$. Moreover, the estimate of alignment is concentrated [6], which means that if we obtain a high alignment on the training set, we can also expect a high alignment on the test set. Thus, high alignment implies good generalization performance.

Fisher Criterion. In this paper, we will use the Fisher quantity [14] for measuring the class separability of the instances data in the kernel feature space induced by kernel function.

Before giving the definition of Fisher, we introduce some notations first. Without loss of generality, we assume that the number of labeled instances in the class, say C_1 (class label equals -1), be n_1 , and the number of labeled instances in the other class, say C_2 (class label equals $+1$), be n_2 , where $n_1 + n_2 = n$. Let $\varphi(\cdot)$ be the kernel feature space induced by K , that is $K(\cdot, \cdot) = \langle \varphi(\cdot), \varphi(\cdot) \rangle$. Let $\bar{\mathbf{t}}, \bar{\mathbf{t}}_1$ and $\bar{\mathbf{t}}_2$ denote the center of entire labeled instances and those of C_1 and C_2 in kernel feature space, that is $\bar{\mathbf{t}} = \frac{1}{n} \sum_{i=1}^n \varphi(\mathbf{x}_i)$, $\bar{\mathbf{t}}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} \varphi(\mathbf{x}_i)^1$, $\bar{\mathbf{t}}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} \varphi(\mathbf{x}_i)^2$, where $\varphi(\mathbf{x}_i)^j$ denotes the i th labeled instance in the j th class ($j = 1, 2$).

Definition 2 (Fisher). Given a kernel function K and the labeled instances $\mathcal{S} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, the Fisher of kernel K with respect to instances \mathcal{S} is the quantity

$$J(\mathcal{S}, K) = (\text{tr}\mathbf{S}_b)/(\text{tr}\mathbf{S}_w), \quad (3)$$

where \mathbf{S}_b is the ‘‘between-class scatter matrix’’, \mathbf{S}_w the ‘‘within-class scatter matrix’’, that is $\mathbf{S}_b = \frac{1}{n} \sum_{i=1}^2 n_i (\bar{\mathbf{t}}_i - \bar{\mathbf{t}})(\bar{\mathbf{t}}_i - \bar{\mathbf{t}})^T$, $\mathbf{S}_w = \frac{1}{n} \sum_{i=1}^2 \sum_{j=1}^{n_i} (\varphi(\mathbf{x}_j)^i - \bar{\mathbf{t}}_i)(\varphi(\mathbf{x}_j)^i - \bar{\mathbf{t}}_i)^T$, ‘‘tr’’ denotes the trace of a matrix.

3 Kernel Construction

In this section, we will describe our new kernel construction method. We will first construct a primal kernel function K_p based on similarity of instances. Then, via conformal transformation [1], we obtain the final kernel K_{Data} . Finally, we determine the coefficients of the kernel function by maximizing alignment and Fisher criteria.

Primal Kernel Function. Without any knowledge of the topology of the space of instances, a learning algorithm is likely to fail. Therefore, on the assumption that there exists an initial inner-product over the input space (we later discuss a way to relax the assumption on the form of the inner-product), we define a primal function as $K_p(\mathbf{x}, \mathbf{z}) = \langle \mathbf{x}, \mathbf{w} \rangle \langle \mathbf{z}, \mathbf{w} \rangle$, where $\mathbf{x}, \mathbf{z}, \mathbf{w} \in \mathcal{X}$.

Note that K_p is symmetrical and for any finite set of points $\{\mathbf{x}_i\}_{i=1}^n$, $\mathbf{x}_i \in \mathcal{X}$, it is easy to verify that the matrix $\mathbf{K}_p = (\mathbf{K}_p(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n$ is positive semi-definite, hence, according to the Proposition [1], K_p is a kernel function.

Obviously, the similarity between two instances \mathbf{x} and \mathbf{z} is high iff both \mathbf{x} and \mathbf{z} are similar (w.r.t the standard inner-product) to the third vector \mathbf{w} .

Analogously, if both \mathbf{x} and \mathbf{z} seem to be dissimilar to the vector \mathbf{w} then they are similar to each other. Thus, $K_p(\mathbf{x}, \mathbf{z})$ can be considered as a measure of the similarity of the instances \mathbf{x} and \mathbf{z} .

Despite the restrictive form of \mathbf{w} , it is still too complex, which may lead to over-fitting. In order to avoid this problem, we further impose two restrictions on \mathbf{w} : 1) \mathbf{w} is restricted to a linear combination of the unlabeled instances $\tilde{\mathcal{S}} = \{\tilde{\mathbf{x}}_i\}_{i=1}^m$; 2) the norm of \mathbf{w} is set to be 1. Thus, K_p can be rewritten as follows

$$K_p(\mathbf{x}, \mathbf{z}) = \langle \mathbf{x}, \mathbf{w} \rangle \langle \mathbf{z}, \mathbf{w} \rangle, \tag{4}$$

where $\mathbf{w} = \sum_{i=1}^m \alpha_i \tilde{\mathbf{x}}_i$ and $\|\mathbf{w}\|^2 = 1$. Here, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)^T$ is the weight coefficient which will be determined later.

Final Kernel Function. Inspired by the paper [11], we use the so-called ‘‘conformal transformation’’ on the primal kernel function K_p as our final kernel function to get better generalization. The final kernel function can be written as

$$K_{\text{Data}}(\mathbf{x}, \mathbf{z}) = q(\mathbf{x})K_p(\mathbf{x}, \mathbf{z})q(\mathbf{z}), \tag{5}$$

where $q(\cdot)$ denotes the factor function, of the form

$$q(\mathbf{x}) = \beta_0 + \sum_{i=1}^l \beta_i K_1(\mathbf{x}, \boldsymbol{\nu}_i), \tag{6}$$

where $K_1(\mathbf{x}, \boldsymbol{\nu}_i) = \exp(-\gamma\|\mathbf{x} - \boldsymbol{\nu}_i\|^2)$. The ‘‘empirical cores’’, denoted by $\{\boldsymbol{\nu}_i, i = 1, 2, \dots, l\}$, can be chosen from the training data or determined according to the distribution of the training data. $\boldsymbol{\beta} = (\beta_1, \dots, \beta_l)^T$ is the combination coefficient which will be discussed later.

Note that K_{data} is symmetrical and it is easy to verify that the matrix $\mathbf{K}_{\text{Data}} = (K_{\text{Data}}(\mathbf{x}_i, \mathbf{x}_j))_{i,j=1}^n$ is positive semi-definite, hence K_{Data} is a kernel function.

In order to determine the weight coefficient $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_m\}^T$ and combination coefficient $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_l\}^T$, we first give two theorems:

Theorem 1. Assume that $\mathbf{M}, \mathbf{N} \in \mathbb{R}^{n \times n}$ are two symmetric matrices and $\boldsymbol{\alpha} \in \mathbb{R}^n$. then, the solution $\boldsymbol{\alpha}^*$ of the following optimization problem

$$\max_{\boldsymbol{\alpha}} \boldsymbol{\alpha}^T \mathbf{M} \boldsymbol{\alpha}, \text{ s.t. } \boldsymbol{\alpha}^T \mathbf{N} \boldsymbol{\alpha} = 1 \tag{7}$$

is the generalized eigenvector corresponding to the largest generalized eigenvalue of the formula $\mathbf{M} \boldsymbol{\alpha}^* = \lambda_{\text{largest}} \mathbf{N} \boldsymbol{\alpha}^*$.

Proof. The dual form of the formula (7) can be written as

$$\max_{\lambda} L(\lambda) = \boldsymbol{\alpha}^T \mathbf{M} \boldsymbol{\alpha} - \lambda(\boldsymbol{\alpha}^T \mathbf{N} \boldsymbol{\alpha} - 1). \tag{8}$$

Let $\frac{\partial L(\lambda)}{\partial \boldsymbol{\alpha}} = 0$, we have

$$\mathbf{M} \boldsymbol{\alpha} - \lambda \mathbf{N} \boldsymbol{\alpha} = 0. \tag{9}$$

Substitute formula (9) to (8), we get

$$\max_{\lambda} L(\lambda) = \lambda.$$

So the solution α^* of the formula (7) is the generalized eigenvector corresponding to the largest generalized eigenvalue of the formula $M\alpha^* = \lambda_{largest}N\alpha^*$.

Theorem 2. Assume that $A, B \in \mathbb{R}^{n \times n}$ are two symmetric matrices and $\beta \in \mathbb{R}^n$. Then the solution β^* of the following optimization problem

$$\max_{\beta} \frac{\beta^T A \beta}{\beta^T B \beta} \tag{10}$$

is the generalized eigenvector corresponding to the largest generalized eigenvalue of the formula $A\beta^* = \lambda_{largest}B\beta^*$.

Proof. The value of the formula (10) is invariable when scaling β , so we can transform formula (10) into the below formula

$$\max_{\beta} \beta^T A \beta, \text{ s.t. } \beta^T B \beta = 1. \tag{11}$$

According to Theorem 1, we know that the solution β^* of the formula (11) is the generalized eigenvector corresponding to the largest generalized eigenvalue of the formula $A\beta^* = \lambda_{largest}B\beta^*$.

Weight Coefficient α . The weight coefficient α will be decided by maximizing the kernel alignment of K_p .

According to the definition of y , we can obtain that $\langle yy^T, yy^T \rangle_F = n^2$, Thus maximizing the kernel alignment $A(\mathcal{S}, K_p)$ can be transformed into the following formula

$$\max \frac{\langle K_p, yy^T \rangle_F}{n\sqrt{\langle K_p, K_p \rangle_F}}, \text{ s.t. } \|w\|^2 = 1. \tag{12}$$

It is not difficult to assert that $\langle K_p, K_p \rangle_F$ is a constant. Thus formula (12) is equivalent to

$$\max \langle K_p, yy^T \rangle_F, \text{ s.t. } \|w\|^2 = 1. \tag{13}$$

Note that $\langle K_p, yy^T \rangle_F = \sum_{r,s=1}^m \alpha_s \alpha_r \sum_{i,j=1}^n y_i y_j \langle x_i, \tilde{x}_r \rangle \langle x_j, \tilde{x}_s \rangle$, and $\|w\|^2 = \|\sum_{r=1}^m \alpha_r \tilde{x}_r\|^2 = \sum_{r,s=1}^m \alpha_r \alpha_s \langle \tilde{x}_r, \tilde{x}_s \rangle$, thus the formula (13) can be written as

$$\max_{\alpha} \alpha^T (C^T B C) \alpha, \text{ s.t. } \alpha^T F \alpha = 1 \tag{14}$$

where matrices $C = (\langle x_i, \tilde{x}_r \rangle)_{i,r=1}^{n,m}$, $B = (y_i y_j)_{i,j=1}^n$ and $F = (\langle \tilde{x}_r, \tilde{x}_s \rangle)_{r,s=1}^m$. According to Theorem 1, the solution α^* of the formula (14) is the generalized eigenvector of $C^T B C \alpha^* = \lambda_{largest} F \alpha^*$. Particularly, if F is a invertible matrix, then α^* is eigenvector corresponding to the largest eigenvalue of $F^{-1} C^T B C$.

Combination Coefficient β . The combination coefficient β will be determined by maximizing Fisher $J(\mathcal{S}, K_{Data})$.

Without loss of generality, let us now assume that the first n_1 labelled instances belong to class -1 , and the remaining n_2 belong to $+1$, $n_1 + n_2 = n$. Then, the kernel matrix of \mathbf{K}_p can be written as $\mathbf{K}_p = \begin{pmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{pmatrix}$.

Let

$$\mathbf{U}_p = \begin{pmatrix} \frac{1}{n_1} \mathbf{K}_{11} & \mathbf{0} \\ \mathbf{0} & \frac{1}{n_2} \mathbf{K}_{22} \end{pmatrix} - \begin{pmatrix} \frac{1}{n} \mathbf{K}_{11} & \frac{1}{n} \mathbf{K}_{12} \\ \frac{1}{n} \mathbf{K}_{21} & \frac{1}{n} \mathbf{K}_{22} \end{pmatrix},$$

$$\mathbf{W}_p = \mathbf{D} - \begin{pmatrix} \frac{1}{n_1} \mathbf{K}_{11} & \mathbf{0} \\ \mathbf{0} & \frac{1}{n_2} \mathbf{K}_{22} \end{pmatrix},$$

where \mathbf{D} is a diagonal matrix whose diagonal elements are $K_p(\mathbf{x}_i, \mathbf{x}_i)$, $i = 1, \dots, n$. Let $\mathbf{q} = (q(\mathbf{x}_1), q(\mathbf{x}_2), \dots, q(\mathbf{x}_n))^T$, it is easy to verify $J(\mathcal{S}, K_{\text{Data}}) = \mathbf{q}^T \mathbf{U}_p \mathbf{q} / (\mathbf{q}^T \mathbf{W}_p \mathbf{q})$, according to formula (6), we have

$$\max_{\boldsymbol{\beta}} J(\mathcal{S}, K_{\text{Data}}) = \max_{\boldsymbol{\beta}} \frac{\boldsymbol{\beta}^T \mathbf{K}_1^T \mathbf{U}_p \mathbf{K}_1 \boldsymbol{\beta}}{\boldsymbol{\beta}^T \mathbf{K}_1^T \mathbf{W}_p \mathbf{K}_1 \boldsymbol{\beta}}, \tag{15}$$

where

$$\mathbf{K}_1 = \begin{pmatrix} 1 & K_1(\mathbf{x}_1, \boldsymbol{\nu}_1) & \dots & K_1(\mathbf{x}_1, \boldsymbol{\nu}_l) \\ 1 & K_1(\mathbf{x}_2, \boldsymbol{\nu}_1) & \dots & K_1(\mathbf{x}_2, \boldsymbol{\nu}_l) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & K_1(\mathbf{x}_n, \boldsymbol{\nu}_1) & \dots & K_1(\mathbf{x}_n, \boldsymbol{\nu}_l) \end{pmatrix}.$$

Let $\mathbf{M}_0 = \mathbf{K}_1^T \mathbf{U}_p \mathbf{K}_1$, $\mathbf{N}_0 = \mathbf{K}_1^T \mathbf{W}_p \mathbf{K}_1$. According to Theorem 2, we know that the solution $\boldsymbol{\beta}^*$ of formula (15) is the generalized eigenvector of $\mathbf{M}_0 \boldsymbol{\beta}^* = \lambda_{largest} \mathbf{N}_0 \boldsymbol{\beta}^*$.

Kernelizing the Kernel. As described above, we assumed that the standard scalar product constitutes the template for the primal kernel function K_p . However, the programming for deciding the weight coefficients and combination coefficients only uses the inner-products of instances, thus we can replace the scalar-product itself with a general kernel function $\kappa : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$, that is replacing $\langle \mathbf{x}, \mathbf{z} \rangle$ with $\kappa(\mathbf{x}, \mathbf{z})$. But using a general kernel function κ , we can not compute the vector \mathbf{w} explicitly. We therefore need to show that the norm of \mathbf{w} for efficiently evaluating the K_{Data} on any two instances.

First note that given the vector $\boldsymbol{\mu}$ we can compute the norm of \mathbf{w} as follows

$$\|\mathbf{w}\|^2 = \left(\sum_r \mu_r \tilde{\mathbf{x}}_r \right)^T \left(\sum_r \mu_s \tilde{\mathbf{x}}_s \right) = \sum_{r,s} \mu_r \mu_s \kappa(\tilde{\mathbf{x}}_r, \tilde{\mathbf{x}}_s).$$

Next, given two vectors \mathbf{x} and \mathbf{z} , we have

$$K_{\text{Data}}(\mathbf{x}, \mathbf{z}) = \sum_{r,s} q(\mathbf{x}_i) q(\mathbf{x}_j) \mu_r \mu_s \kappa(\mathbf{x}, \tilde{\mathbf{x}}_r) \kappa(\mathbf{z}, \tilde{\mathbf{x}}_s).$$

The pseudo code of kernel construction is given in Algorithm [1](#).

Input: $\mathcal{S} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$, $\tilde{\mathcal{S}} = \{\tilde{\mathbf{x}}_i\}_{i=1}^m$

Output: α^*, β^*

- 1 Calculate matrix \mathbf{B} , \mathbf{C} , \mathbf{F}
- 2 Find the generalized eigenvector α^* for the problem $\mathbf{C}^T \mathbf{B} \mathbf{C} \alpha^* = \lambda_{largest} \mathbf{F} \alpha^*$
- 3 Calculate $\mathbf{w} = \sum_{i=1}^m \alpha_i^* \tilde{\mathbf{x}}_i / \|\sum_{i=1}^m \alpha_i^* \tilde{\mathbf{x}}_i\|$
- 4 Calculate matrix \mathbf{M}_0 , \mathbf{N}_0
- 5 Find the generalized eigenvector β^* for the problem $\mathbf{M}_0 \beta^* = \lambda_{largest} \mathbf{N}_0 \beta^*$
- 6 Return: α^*, β^*

Algorithm 1. Kernel Construction Algorithm

4 Experimental Results and Analysis

In this section, we will verify the feasibility and robustness of our new method.

We perform experiments on platform MATLAB 7.0.1(R14) Service Pack 1. The standard datasets are chosen from UCI, Statlog and Delve. The detail of datasets are shown in Table [1](#).

We randomly split the datasets into training data and test data (test data is also regarded as unlabeled instances). We choose a random set of one-third of the data points as the empirical cores $\{\nu_i, i = 1, \dots, l\}$. All the experiments are conducted by adapting L-1 Support Vector Machine (SVM) [\[5\]](#). To

Table 1. Datasets used in Experiments. The table shows the number of features and the number of training and testing instances.

Dataset	Size	Dimension	Training Size	Test Size
Sonar	208	60	69	70
Heart	270	13	90	90
Liver	345	6	115	115
Ionosphere	351	34	117	117
Breast	683	10	227	228
Australian	690	14	230	230
Diabetes	768	8	256	256
Fourclass	862	2	287	288
German	1000	24	333	334
Titanic	2201	3	733	734

Table 2. Optimal classification test accuracies of K_{Gauss} , K_{Poly} , K_{Boosting} and K_{Data} on ten datasets (%)

	Sonar	Heart	Liver	Ionosphere	Breast	Australian	Diabetes	Fourclass	German	Titanic
K_{Gauss}	79.6	79.3	72.1	94.3	96.6	84.8	76.2	99.9	71.7	78.9
K_{Poly}	78.1	82.5	67.8	89.4	96.8	85.5	75.9	99.8	75.9	79.5
K_{Boosting}	79.2	80.3	71.5	84.8	95.8	85.3	76.5	94.4	75.1	80.4
K_{Data}	80.2	84.3	76.8	85.2	97.1	85.8	76.1	93.6	76.3	80.9

Table 3. Classification test accuracies with different values of penalty factor C on Heart

C	K_{Gauss}	K_{Poly}	K_{Boosting}	K_{Data}
10^0	76.3	81.5	80.3	84.3
10^1	77.4	84.5	76.8	84.9
10^2	79.3	81.3	77.2	84.1
10^3	77.7	83.7	81.3	84.2
10^4	78.8	84.0	75.6	84.8

Table 4. Classification test accuracies with different value of penalty factor C on Liver

C	K_{Gauss}	K_{Poly}	K_{Boosting}	K_{Data}
10^0	67.2	62.3	71.5	76.8
10^1	68.1	68.1	72.5	75.3
10^2	69.5	70.1	74.5	76.8
10^3	66.9	67.7	72.1	77.2
10^4	61.4	68.1	75.6	76.5

reduce statistical variability, results here are based on averages over 50 random repetitions.

Feasibility. We will verify the feasibility of our method via contrasting the kernel function obtained by our new method to other three widely used kernel functions on classification test accuracy. The denotations of four kernel functions are given in the following:

- Gauss-kernel function $K_{\text{Gauss}}(\mathbf{x}, \mathbf{z}) = \exp(-\gamma_0 \|\mathbf{x} - \mathbf{z}\|^2)$,
- Poly-kernel function $K_{\text{Poly}}(\mathbf{x}, \mathbf{z}) = (1 + \langle \mathbf{x}, \mathbf{z} \rangle)^d$,
- kernel function constructed by the state-of-art kernel construction method [4], K_{Boosting} ,
- kernel function constructed by our new method, K_{Data} .

The Gauss kernel parameter $\gamma_0 \in \{0.5, 1, 2, 10\}$, Poly-kernel parameter $d \in \{1, 2, 10, 100\}$. Set the parameter γ of K_{Data} and the penalty factor C of SVM to 1. In the next subsection of robustness experiments, we will consider the fluctuation of classification test accuracy when we change the values of γ and C .

Table 5. Test accuracies of K_{Data} with different value of γ on ten datasets (%)

γ	Sonar	Heart	Liver	Ionosphere	Breast	Australian	Diabetes	Fourclass	German	Titanic
0.5	79.6	83.3	77.0	84.9	96.0	85.0	76.1	92.5	75.3	80.5
1.0	80.2	84.3	76.8	85.2	97.2	86.5	76.1	93.6	76.5	81.3
1.5	79.7	84.2	77.2	85.1	96.8	87.0	77.0	93.5	77.0	82.2
2.0	80.1	84.3	77.5	85.0	97.2	85.0	75.0	95.4	77.8	83.4
2.5	80.7	84.5	77.0	85.7	96.8	86.2	74.2	95.4	75.3	81.3
3.0	81.2	84.1	77.2	84.0	96.6	85.2	75.1	95.1	76.0	82.1

The optimal parameters are obtained via 10 fold cross-validation on training set. We then account classification test accuracy of each kernel function with optimal parameters on test data, with results shown in Table 2.

We can see that K_{Data} is outstanding with the optimal accuracy in seven datasets of ten, and with accuracy closing to the optimal in others. Thus, it can show that our kernel construction method is feasible.

Robustness. We consider the fluctuation of classification test accuracy of K_{Data} when we change the values of penalty factor C and γ . $C \in \{10^0, 10^1, 10^2, 10^3, 10^4\}$ and $\gamma = 1$, the rest parameters are the same as in above subsection experiment. The corresponding classification test accuracies of the four kernel functions on Heart and Liver are reported in Table 3 and Table 4. The similar results can be obtained in other datasets. We can see that K_{Data} is stabler than other kernels respecting to C .

Then, we set $\gamma \in \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$ and $C = 1$. In Table 5, the classification test accuracies of K_{Data} on ten datasets are reported, the maximum deviation is only 2% on Australian and Diabetes which reflects the insensitivity with respect to γ .

By means of the above experiments, we know that the kernel constructed by our method is robust with respect to C and γ , which is important for the application of the method.

5 Conclusion

In this paper, we present a method for constructing a kernel function by simply solving a generalized eigenvalue problem. Comparing with existing kernel construction methods, our method can avoid the local maxima problem, and the kernel constructed by our method is positive semidefinite. In addition, experimental results on the benchmark datasets show that, our method is effective and robust.

We will focus on the explicit description of the reproducing kernel Hilbert space of the kernel function constructed by our method in the future.

Acknowledgements. This work was supported by Natural Science Foundation of China under Grant No.61170019 and Natural Science Foundation of Tianjin under Grant No. 11JCYBJC00700.

References

1. Amari, S., Wu, S.: Improving support vector machine classifiers by modifying kernel functions. *IEEE Transactions on Neural Networks* 12(6), 783–789 (1999)
2. Aronszajn, N.: Theory of reproducing kernels. *Transactions of the American Mathematical Society* 68(3), 337–404 (1950)
3. Chapelle, O., Vapnik, V.: Model selection for support vector machines. In: *Advances in Neural Information Processing Systems* 12, pp. 230–236. MIT Press, Cambridge (1999)

4. Crammer, K., Keshet, J., Singer, Y.: Kernel design using boosting. In: *Advances in Neural Information Processing Systems 15*, pp. 537–544. MIT Press, Cambridge (2003)
5. Cristianini, N., Shawe-Taylor, J.: *An introduction to support Vector Machines: and other kernel-based learning methods*. Cambridge University Press (2000)
6. Cristianini, N., Shawe-taylor, J., Elisseeff, A., Kandola, J.: On kernel-target alignment. In: *Advances in Neural Information Processing Systems 14*, pp. 367–373. MIT Press, Cambridge (2002)
7. Hertz, T., Hillel, A., Weinshall, D.: Learning a kernel function for classification with small training samples. In: *Proceedings of the 23rd international Conference on Machine Learning*, pp. 401–408. ACM Press, New York (2006)
8. Merler, S., Jurman, G., Furlanello, C.: Deriving the Kernel from Training Data. In: Haindl, M., Kittler, J., Roli, F. (eds.) *MCS 2007*. LNCS, vol. 4472, pp. 32–41. Springer, Heidelberg (2007)
9. Mika, S., Ratsch, G., Weston, J., Scholkopf, B., Mullers, K.: Fisher discriminant analysis with kernels. In: *Proceedings of the 1999 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing IX*, pp. 41–48. IEEE (1999)
10. Ong, C., Smola, A., Williamson, R.: Hyperkernels. In: *Advances in Neural Information Processing Systems 15*, pp. 478–485. MIT Press (2003)
11. Scholkopf, B., Smola, A., Muller, K.: Kernel Principal Component Analysis. In: Gerstner, W., Hasler, M., Germond, A., Nicoud, J.-D. (eds.) *ICANN 1997*. LNCS, vol. 1327, pp. 583–588. Springer, Heidelberg (1997)
12. Sollich, P.: Bayesian methods for support vector machines: Evidence and predictive class probabilities. *Machine Learning* 46(1), 21–52 (2002)
13. Vapnik, V.: *The nature of statistical learning theory*. Springer, Berlin (2000)
14. Xiong, H., Swamy, M., Ahmad, M.: Optimizing the kernel in the empirical feature space. *IEEE Transactions on Neural Networks* 16(2), 460–474 (2005)

Probabilistic Model Combination for Support Vector Machine Using Positive-Definite Kernel-Based Regularization Path

Ning Zhao, Zhihui Zhao, and Shizhong Liao

School of Computer Science and Technology, Tianjin University, Tianjin, China
szliao@tju.edu.cn

Abstract. Model combination is an important approach to improving the generalization performance of support vector machine (SVM), but usually has low computational efficiency. In this paper, we propose a novel probabilistic model combination method for support vector machine on regularization path (PM-CRP). We first design an efficient regularization path algorithm, namely the regularization path of support vector machine based on positive-definite kernel (PDSVMP), which constructs the initial candidate model set. Then, we combine the initial models using Bayesian model averaging. Experimental results on benchmark datasets show that PMCRP has significant advantage over cross-validation and the Generalized Approximate Cross-Validation (GACV), meanwhile guaranteeing high computation efficiency of model combination.

Keywords: Support Vector Machines, Model Combination, Regularization Path.

1 Introduction

Support vector machine (SVM, [8]) that emerged in recent years has been a widely used tool for classification. The generalization performance of SVM depends on the parameters of regularization and kernels. Various algorithms [2], [1] have been developed for choosing the best parameters. However, single model only has limited information and usually exists uncertainty. Model combination can integrate all useful information from the candidate models into the final hypothesis to guarantee good generalization performance.

In this paper, we propose a novel probabilistic model combination method for support vector machine on regularization path. First, we develop an efficient regularization path algorithm PDSVMP. Then, the probabilistic interpretation of support vector machine is described. The SVM prior is simply a Gaussian process (GP) over the models, while model posterior probability is obtained by Bayesian formula, and the combination on the initial models is achieved using Bayesian model averaging. Finally, we conduct several experiments to test the performance of the new method.

The rest of the paper is organized as follows. Section 2 discusses the support vector machine and its probabilistic interpretation. In section 3, we develop an efficient regularization path algorithm PDSVMP. In section 4, we present our model combination method PMCRP. The experimental results are given in section 5. Section 6 discusses the outline perspectives for future work and concludes this paper.

2 Preliminaries

Suppose we have a dataset $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$, where $x_i \in \mathbf{X} \subset \mathbb{R}^p$ is an input vector, and binary outputs $y_i = \pm 1$ corresponding to the two classes.

Suppose $f = h + b \in (\mathcal{H}_k + \mathbb{R})$, where \mathcal{H}_k is a reproducing kernel Hilbert space (RKHS). The decision function $f(x)$ generated by a L1-SVM corresponds to the solution of

$$\min_{f \in \mathcal{H}_k} \left\{ \sum_{i=1}^n L(y_i, f(x_i)) + \frac{\lambda}{2} \|h\|_{\mathcal{H}_k}^2 \right\}, \quad (1)$$

where $\lambda > 0$ is the regularization parameter and $\|\cdot\|_{\mathcal{H}_k}$ is the RKHS norm. Loss function $L(a)$ is the ‘‘hinge loss’’, while we use a shifted form $L(a) = (1 - a)H(1 - a)$, $H(1 - a)$ is the Heaviside step function ensures that this is zero for $a > 1$. The SVM decision function f takes the form $f(x) = \sum_{j=1}^n \mathbf{w} \cdot \phi(x_j) + b$. Find \mathbf{w} and b to minimize the following SVM optimization problem:

$$\sum_{i=1}^n L(y_i f(x_i)) + \frac{\lambda}{2} \|\mathbf{w}\|^2. \quad (2)$$

The first term in (2) can be written as the negative log of probability of obtaining output y for a given x and f , $\Pr(y|x, f) = \kappa \exp[-L(yf(x))]$, and the constant κ is chosen to normalize $\Pr(y|x, f)$ across all possible values of y . The second term in (2) gives the prior of \mathbf{w} . The SVM prior is therefore simply a Gaussian Process (GP) over the functions f , with zero mean, and its covariance function is $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) + v^2$. More details can be found in [7].

3 PDSVMP Algorithm

Hastie et al. [3] derived an algorithm SVMpath that could fit the entire path of SVM solutions for every value of λ , with essentially the same computational cost as fitting one SVM model. Ong et al. [4] proposed an improved algorithm based on SVMpath, named ISVMP. ISVMP could correctly handle those data sets contain duplicate samples, nearly duplicate samples or linear dependent samples, but also has a high computational price.

One key step of the regularization path algorithm is to solve a linear equation system [3]:

$$A_l * \delta^a = \mathbf{1}^a, \quad (3)$$

$$A_l = \begin{pmatrix} 0 & y_l^T \\ y_l & K_l^* \end{pmatrix}, \delta^a = \begin{pmatrix} \delta_0 \\ \delta \end{pmatrix}, \mathbf{1}^a = \begin{pmatrix} 0 \\ \mathbf{1} \end{pmatrix}.$$

SVMpath assumes that A_l has full rank, then we can write $b^a = A_l^{-1} \mathbf{1}^a$, where $b^a = \frac{\delta^a}{\lambda_l - \lambda}$. Unfortunately the matrix A_l is not positive definite, we develop a new method to solve this linear equation system. The second row of (3) can be re-written as $K_l^* (K_l^{*-1} y_l b_0 + b) = \mathbf{1}$. Then $y_l^T K_l^{*-1} \mathbf{1} = y_l^T K_l^{*-1} y_l b_0$. The system of linear equations can then be re-written as

$$\begin{pmatrix} y_l^T K_l^{*-1} y_l & \mathbf{0}^T \\ \mathbf{0} & K_l^* \end{pmatrix} \begin{pmatrix} b_0 \\ b + K_l^{*-1} y_l b_0 \end{pmatrix} = \begin{pmatrix} y_l^T K_l^{*-1} \mathbf{1} \\ \mathbf{1} \end{pmatrix}. \quad (4)$$

Suppose the kernel matrix is positive-definite. In this case, the matrix on the left hand side is positive-definite, as K_l^* is positive-definite, and $y_l^T K_l^{*-1} y_l$ is positive since the inverse of a positive definite matrix is also positive definite. Hence the system of linear equations has a unique solution.

Another key step of the regularization path algorithm is to compute the next value of the regularization parameter λ_{l+1} . For $\lambda_{l+1} < \lambda < \lambda_l$, let $\delta_\lambda = \lambda - \lambda_l$, then we get the maximum

$$\delta_\lambda^* = \max\{\delta_\lambda^{\mathcal{L}}, \delta_\lambda^{\mathcal{R}}, -\lambda_l\}. \tag{5}$$

Where $\delta_\lambda^{\mathcal{L}} = \max_{i \in \mathcal{L}} \{ \{1 - \alpha_i^l / b_a^i : b_a^i < 0\}, \{-\alpha_i^l / b_a^i : b_a^i > 0\} \}$, $\delta_\lambda^{\mathcal{R}} = \max_{i \in \mathcal{R}} \{ \lambda_l \xi_i^l / s_i : s_i > 0 \}$, $\delta_\lambda^{\mathcal{L}} = \max_{i \in \mathcal{L}} \{ \lambda_l \xi_i^l / s_i : s_i < 0 \}$. $s_i = b_a^T A_i^* - 1, i \in \mathcal{L} \cup \mathcal{R}$ and ξ_i is the slack variance of the soft margin SVM. In this case, $\lambda^{l+1} = \delta_\lambda^* + \lambda^l$. Expression (5) is clearly different from those used in SVMpath or ISVMP for determining the next event.

4 Model Combination for Support Vector Machine

4.1 Initial Model Set

PDSVMP can fit the entire path of SVM solutions for every value of the regularization parameter. However, the pursued policy can have a high computational price, the family of piecewise linear solution path is of particular interest.

Definition 1 (Initial model set). The initial candidate model set \mathcal{M} we considered corresponding to the functions of all break points on the regularization path, i.e. $\mathcal{M} = \{f_{\lambda_l} | \lambda_l \in [0, +\infty), l = 1, \dots, N\}$. N is the iteration number of the PDSVMP algorithm.

Model Posterior Probability. We choose a Gaussian Process prior on f_{λ_l} . The link between SVM and Gaussian process (GPs) has been pointed out by many authors, e.g. [5]. The GP prior $\Pr(f_\lambda)$ with covariance function equals to the kernel $K(x_i, x_j)$

$$\Pr(f_\lambda) = \frac{1}{\sqrt{\det(2\pi K)}} \exp \left\{ -\frac{1}{2} \sum_{i,j} f_\lambda(x_i) K^{-1}(x_i, x_j) f_\lambda(x_j) \right\}. \tag{6}$$

We simply compute the model posterior by combining the frequentist-motivated SVM with a Bayesian interpretation. Due to the point of normalized, the SVM loss is unnatural to interpret as the negative log of probability of the output y . We normalize the SVM loss pointwise similar to [7] and [6], i.e. let $\kappa = \exp(-L(f_\lambda(x))) + \exp(-L(-f_\lambda(x)))$. For each model f_λ in \mathcal{M}

$$\Pr(y|x, f_\lambda) = \begin{cases} \frac{1}{1 + \exp\{-2yf_\lambda(x)\}} & \text{for } |f_\lambda(x)| \leq 1, \\ \frac{1}{1 + \exp\{-y[f_\lambda(x) + \text{sgn}(f_\lambda(x))]\}} & \text{others.} \end{cases} \tag{7}$$

According to Bayes' formula, the model posterior of f_λ is $\Pr(f_\lambda | x, y) = \frac{\Pr(f_\lambda) \Pr(y|x, f_\lambda)}{\Pr(y|x)}$.

4.2 Combination of Support Vector Machine Probabilistic Models

Suppose the test dataset is $T = \{(x_1, y_1), \dots, (x_s, y_s)\}$, our main interest is to predict the label of new input vector $x_* \in T$, the Bayesian model averaging over \mathcal{M} is

$$\Pr(y_*|x_*, D) = \sum_{f_{\lambda_l} \in \mathcal{M}} \Pr(y_*|x_*, f_{\lambda_l}, D) \Pr(f_{\lambda_l}|D). \quad (8)$$

For each model in \mathcal{M} , the posterior probability $\Pr(y_*|x_*, f_{\lambda})$ can be computed from (7), and the model posterior $\Pr(f_{\lambda_l}|D)$ can be computed using Bayes' formula.

Computational Complexity. The main computational burden of PMCRP centers on the solution of problem (1). Letting m be the average size of \mathcal{E}_l , then the approximate computational cost of SVMpath is $O(cn^2m + nm^2)$. In addition, the computation of $\Pr(f_{\lambda}|x, y)$ is $O(N)$, i.e. $O(cn)$. Suppose the size of test dataset T is s , then computing $\Pr(y_*|x_*, f(x_*))$ and y'_* are both $O(s)$. The total computational complexity of PMCRP is hence $O(cn^2m + nm^2 + s)$, where s is usually far less than the others.

5 Experiments

Experiment Setup and Data. In this section, we conduct several experiments to evaluate the performance of our proposed PMCRP. Table 1 gives a summary of these datasets. We choose the radial basis function (RBF) kernel with a single parameter, that is $K(x, x') = \exp(-\gamma||x - x'||^2)$, where γ is the kernel parameter, and the range of γ is $\{0.001, 0.01, 0.2, 0.8, 1, 5, 10, 50, 100, 200\}$. We compare PMCRP with alternatives based on estimates of the test error, namely 3-fold cross-validation (CV) and Wahba's GACV [9].

Table 1. Summary of the Benchmark Datasets: n —the size of the training data; m —the size of the test data; p —the dimension of the input vector

Datasets	n	m	p	Datasets	n	m	p
Sonar	108	100	60	Heart	162	108	13
Liver	207	138	6	Ionosphere	210	141	34
Breast	410	274	10	Australian	414	276	14
Diabetes	460	308	8	Fourclass	517	345	2
German	600	400	24	Splice	600	400	60

The Feasibility and the Efficiency. We first examine the feasibility of PMCRP. In order to reduce the impact of kernel parameters, all three methods are taken over all values of γ within its range, and select the lowest average test error, respectively. The results are shown in table 2.

From the table we can find that the model combination method has the lowest prediction error against over other methods in most datasets, except for datasets diabetes and splice. In a sense, this experiment shows that PMCRP has a good generalization

Table 2. The rate of test errors for various methods

Datasets	3-fold CV	GACV	PMCRP
Sonar	0.13095240	0.09523810	0.01190476
Heart	0.13888900	0.19444400	0.05555600
Liver	0.28985510	0.23913040	0.14492750
Ionosphere	0.06382979	0.04964539	0.02836879
Breast	0.03284700	0.04014600	0.02554745
Australian	0.11956500	0.14130430	0.00362319
Diabetes	0.23052000	0.25324700	0.28571400
Fourclass	0.00579710	0.00869565	0.00289855
German	0.21750000	0.27500000	0.02500000
Splice	0.15333333	0.16666667	0.22250000

Table 3. Run time comparison (Seconds)

Datasets	GACV	PMCRP	Datasets	GACV	PMCRP
Sonar	0.20	0.18	Heart	1.46	1.31
Liver	3.62	2.95	Ionosphere	1.01	0.79
Breast	12.83	8.99	Australian	24.10	19.89
Diabetes	18.08	12.68	Fourclass	40.90	26.65
German	30.10	20.28	Splice	71.55	37.66

performance. We further verify the efficiency of PMCRP. The results are shown in Table 3. From the tables we can find that the time-consuming of PMCRP is less than that of the GACV.

6 Conclusions and Further Research

In this paper, we have proposed the novel model combination method, called PMCRP, for SVM on regularization path, which has a computational complexity of $O(cn^2m + nm^2 + s)$. The regularization path algorithm based on positive-definite kernel is more practical than the original SVMpath, and the Bayesian model averaging is feasible. Experimental results on benchmark datasets verify the feasibility and the efficiency of PMCRP.

Further research aims to extend our framework to support vector regression, and give an efficient combination method for regression models.

Acknowledgements. This work was supported by Natural Science Foundation of China under Grant No.61170019 and Natural Science Foundation of Tianjin under Grant No.11JCYBJC00700.

References

1. Adankon, M.M., Cheriet, M.: Optimizing resources in model selection for support vector machine. *Pattern Recognition* 40(3), 953–963 (2007)
2. Chapelle, O., Vapnik, V., Bousquet, O., Mukherjee, S.: Choosing multiple parameters for support vector machines. *Machine Learning* 46(1), 131–159 (2002)
3. Hastie, T., Rosset, S., Tibshirani, R., Zhu, J.: The entire regularization path for the support vector machine. *Journal of Machine Learning Research* 5, 1391–1415 (2004)
4. Ong, C.J., Shao, S., Yang, J.: An improved algorithm for the solution of the regularization path of support vector machine. *IEEE Transactions on Neural Networks* 21(3), 451–462 (2010)
5. Opper, M., Winther, O.: Gaussian process classification and svm: mean field results and leave-one-out estimator. In: Smola, A.J., Bartlett, P., Schölkopf, B., Schuurmans, D. (eds.) *Advances in Large Margin Classifiers*, pp. 43–65. MIT Press, Cambridge (2000)
6. Seeger, M.: Bayesian model selection for support vector machine, gaussian processes and other kernel classifiers. In: *Advances in Neural Information Processing Systems 12*, pp. 603–609. MIT Press, Cambridge (2000)
7. Sollich, P.: Bayesian methods for support vector machines: evidence and predictive class probabilities. *Machine Learning* 46(1-3), 21–52 (2002)
8. Vapnik, V.N.: *The nature of statistical learning theory*. Springer, New York (2000)
9. Wahba, G., Lin, Y., Zhang, H.: Generalized approximate cross validation for support vector machines, or, another way to look at margin-like quantities. Tech. rep., Department of Statistics, University of Wisconsin (1999)

The Relationship of Filters in Lattice Implication Algebra

Ruijuan Lv and Yang Xu

Intelligent Control Development Center, Southwest Jiaotong University,
Chengdu 610031, China

lvrujuan2008@gmail.com, xuyang@home.swjtu.edu.cn

Abstract. In this paper, we focus on the properties of filters in lattice implication algebra. We study the relationship of associative filter and implicative filter, n -fold associative filter and n -fold implicative filter in detail. And a sufficient condition of involution filter in lattice implication algebra is proved. Then the relationship of some filters is given in Figure 4, and the transformation conditions among these filters are analyzed and obtained in Figure 5. Last, some properties of filter lattice are discussed.

Keywords: Lattice implication algebra, Filter.

1 Introduction

In order to study lattice valued logic system more deeply, Xu[1] proposed the concept of lattice implication algebras. Since it has vital potential application, up to now, they have been widely studied. In particularly, emphasis seems to have been put on the ideals and filters theory. In 1993, the properties and relations of the filters and implicative filters were systematically discussed in lattice implication algebra by Xu and Qin [6]. In 1998, Jun, Xu and Qin proposed the concepts of the positive implicative filter and associative filter in lattice implication algebra [7]. In 1998, Meng proposed the concept of the prime filter in lattice implication algebra, and some properties of it were given [8]. In 1999, Qin and Xu presented the concept of the ultra-filter in lattice implication algebra [9]. In 2001, Wang, Xu and Song proposed the concept of the obstinate filter in lattice implication algebra, and proved that obstinate filter, maximal positive implicative filter and maximal implicative filter are equivalent [10]. In 2004, n -fold positive implicative filter of lattice implication algebra was introduced and the related Song and Jun [12]. In 2010, Wu introduced the concept of Boolean filter in lattice implication algebra, investigated the related properties and proved that Boolean filter and implicative filter were equivalent in [4]. The concept of submaximal filter w.r.t x in lattice implication algebra and its properties were investigated by Du [5]. At the same time, the relation of implication filter, prime filter and ultra filter was discussed in [5].

This present paper mainly focused on the properties of filters, the relationship of filters, and into each other's conditions of filters in lattice implication algebra. The structure of this paper is organized as follows. In section 2, some basic concepts and properties in lattice implication algebra are recalled. In section 3, the properties of filters in lattice implication algebra are given. The relation of some filters is shown in

Figure 3.1. In addition, the transformation conditions among these filters are obtained in Figure 3.2. In section 4, filter lattice and its properties are also discussed in lattice implication algebra.

2 Preliminaries

In this section, we recall basic concepts and some properties in lattice implication algebra.

Definition 2.1[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a bounded lattice with an order-reversing involution $'$, I and O the greatest and the smallest element of L , respectively. And

$$\rightarrow: L \times L \rightarrow L$$

is a mapping and satisfies the following conditions for any $x, y, z \in L$:

- $(I_1) x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$;
- $(I_2) x \rightarrow x = I$;
- $(I_3) x \rightarrow y = y' \rightarrow x'$;
- $(I_4) x \rightarrow y = y \rightarrow x = I$ implies $x = y$;
- $(I_5) (x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x$;
- $(L_1) (x \vee y) \rightarrow z = (x \rightarrow z) \wedge (y \rightarrow z)$;
- $(L_2) (x \wedge y) \rightarrow z = (x \rightarrow z) \vee (y \rightarrow z)$.

Then $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ is called a lattice implication algebra.

Example 2.1[1]. Let $L = \{O, a, b, c, d, I\}$, $O' = I, a' = c, b' = d, c' = a, d' = b, I' = O$, The Hasse diagram be defined as Fig.1, and operator of L be defined as Table 1, then $(L, \vee, \wedge, ', \rightarrow)$ is a lattice implication algebra.

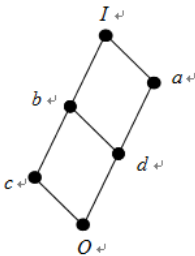


Fig. 1. Hasse diagram of L_6

Table 1. Implication Operator of L_6

\rightarrow	O	a	b	c	d	I
O	I	I	I	I	I	I
a	c	I	b	c	b	I
b	d	a	I	b	a	I
c	a	a	I	I	a	I
d	b	I	I	b	I	I
I	O	a	b	c	d	I

Example 2.2. Let $L = \{O, a, b, I\}$, the partial order on L be defined as $O < a < b < I$, and $O' = I, a' = b, b' = a, I' = O$, The Hasse diagram be defined as Fig.2, and operator of L be defined as Table 2 Then $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ is a lattice implication algebra.

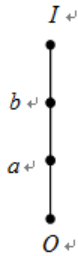


Fig. 2. Hasse diagram of L_4

Table 2. Implication Operator of L_4

\rightarrow	O	a	b	I
O	I	I	I	I
a	b	I	I	I
b	a	b	I	I
I	O	a	b	I

Example 2.3[1]. Let $L = \{O, a, b, I\}$ and $O' = I, a' = b, b' = a, I' = O$, The Hasse diagram be defined as Fig.3, and operator of L be defined as Table 3, then $(L, \vee, \wedge, ', \rightarrow, O, I)$ is a lattice implication algebra.

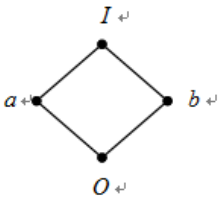


Fig. 3. Hasse diagram of L_4

Table 3. Implication Operator of L_4

\rightarrow	O	a	b	I
O	I	I	I	I
a	b	I	b	I
b	a	a	I	I
I	O	a	b	I

Definition 2.2[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $J \subseteq L$ is said to be a filter of \mathcal{L} , if it satisfies the following conditions:

- (1) $I \in J$;
- (2) For any $x, y \in L$, if $x \in J$ and $x \rightarrow y \in J$, then $y \in J$.

Definition 2.3[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, J is said to be a implicative filter of \mathcal{L} , if it satisfies the following conditions:

- (1) $I \in J$;
- (2) for any $x, y, z \in L$, if $x \rightarrow (y \rightarrow z) \in J$ and $x \rightarrow y \in J$, then $x \rightarrow z \in J$.

Definition 2.4[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. A subset F of L is called a positive implicative filter of \mathcal{L} if it satisfies:

- (1) $I \in F$;
- (2) $x \rightarrow ((y \rightarrow z) \rightarrow y) \in F$ and $x \in F$ imply $y \in F$ for any $x, y, z \in F$.

Definition 2.5[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $F \subseteq L$, $F \neq \emptyset$, $x \in L$, for any $y, z \in L$, F is called an associative filter of \mathcal{L} w.r.t x if it satisfies:

- (1) $I \in F$;
- (2) If $x \rightarrow (y \rightarrow z) \in F$, and $x \rightarrow y \in F$, then $z \in F$.

Definition 2.6[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, P is a proper filter of \mathcal{L} . P is called a prime filter if and only if $a \vee b \in P$ implies $a \in P$ or $b \in P$ for any $a, b \in L$.

Theorem 2.1[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, P is a proper filter of \mathcal{L} . P is a prime filter if and only if $a \rightarrow b \in P$ or $b \rightarrow a \in P$ for any $a, b \in L$.

Definition 2.7[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. A filter J of \mathcal{L} is called an ultra-filter if for any $x \in L$, $x \in J$ if and only if $x' \notin J$.

Theorem 2.2[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. F is a proper filter of \mathcal{L} . F is an ultra-filter if and only if $a \oplus b \in F$ implies $a \in F$ or $b \in F$ for any $a, b \in L$.

Definition 2.8[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. J is a filter of \mathcal{L} . J is called an obstinate filter if $x \notin J, y \notin J$ imply $x \rightarrow y \in J, y \rightarrow x \in J$ for any $x, y \in L$.

Theorem 2.3[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. If J is an ultra-filter, then J is an obstinate filter.

Definition 2.9[2]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. $F \subseteq L, F \neq \emptyset, n \in N$, for any $x, y, z \in L$, F is called an n -fold implicative filter of \mathcal{L} if it satisfies:

- (1) $I \in F$;
- (2) If $x^n \rightarrow (y \rightarrow z) \in F$, and $x^n \rightarrow y \in F$, then $x^n \rightarrow z \in F$.

Definition 2.10. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $F \subseteq L$, $F \neq \emptyset$, $n \in \mathbb{N}$, $x \in L$, for any $y, z \in L$, F is called a n -fold associative filter of \mathcal{L} w.r.t x if it satisfies:

- (1) $I \in F$;
- (2) If $x^n \rightarrow (y \rightarrow z) \in F$, and $x^n \rightarrow y \in F$, then $z \in F$.

Theorem 2.4[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, F is a filter, if for any $x, y \in L$, $x \rightarrow (x \rightarrow y) \in F$ imply $x \rightarrow y \in F$, then F is an implicative filter.

Definition 2.11[4]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, J is a filter, $J \subseteq L$ is said to be a Boolean filter if $x \vee x' \in J$ for any $x \in L$.

Definition 2.12[5]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, F is a proper filter of \mathcal{L} , for any $x \in L$, F is called a submaximal filter w.r.t x , if F is the maximal filter of the filters not contained x , i.e., if F_1 is a filter of L and $F \subseteq F_1$, then $F = F_1$ or $x \in F_1$.

Theorem 2.5[5]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ is a lattice implication algebra, for any $x \in L$, and $x \neq I$, then there exist a submaximal filter w.r.t x .

Definition 2.13[1]. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $a \in L$. F_a is called the I-filter determined by a if $F_a = \{x \mid x \in L, x \vee a = I\}$.

For any $A \subseteq L$, we denote A^* the set $\{x \mid x \vee a = I \text{ for any } a \in A\}$. It is clear that $A^* = \bigcap_{a \in A} F_a$ and A^* is a filter.

Definition 2.14[1]. A filter F of \mathcal{L} is called an involution filter if $F = (F^*)^*$.

3 The Relationship of Filters in Lattice Implication Algebra

Denote $a \in L$, $A(a) = \{x \mid x \in L, a \leq x\}$. $F(\mathcal{L})$ is the set of filters in L .

Theorem 3.1. Let $\mathcal{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $F \in F(\mathcal{L})$, F is an associative filter w.r.t x , for any $x \in F$, $x \neq O$.

Proof. For any $x \in F$, $x \neq O$. Suppose $x \rightarrow (y \rightarrow z) \in F$ and $x \rightarrow y \in F$, $x \in F \in F(\mathcal{L})$, it follows that $y \rightarrow z \in F$ and $y \in F \in F(\mathcal{L})$, then $z \in F$, F is an associative filter w.r.t x .

Remark 3.1. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $F \subseteq L, F$ is an associative filter w.r.t $x, x \in L, F$ may not be a filter of \mathcal{L} .

In Example 2.2, $\{O, a, I\}$ is an associative filter w.r.t a , but it is not a filter of \mathcal{L}_4 .

Theorem 3.2. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, for any $F \in F(\mathcal{L})$, if $A(a)$ is a filter for any $a \in L, F$ is an implicative filter.

Proof. Suppose that $A(a)$ is a filter for any $a \in L$, For any $x, y, z \in L, A(x)$ is a filter for any $x \in L, F \subseteq F(\mathcal{L})$, then $F \subseteq \{A(x) \mid A(x) \text{ is an filter for any } x \in L\}$, if $x \rightarrow (y \rightarrow z) \in F$ and $x \rightarrow y \in F$, then $y \rightarrow z \in A(x)$ and $y \in A(x), z \in A(x)$, it follows $z \geq x, x \rightarrow z = I \in F$, so F is an implicative filter of L .

But the converse is not always true. In Example 2.1, the filter $\{c, b, I\}$ is an implicative filter, but $A(b) = \{b, I\}$ is not a filter.

Theorem 3.3. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $F \subseteq L, F$ is an associative filter w.r.t x for any $x \neq O$, then F is an implicative filter.

Proof. Suppose $x \rightarrow (y \rightarrow z) \in F$ and $x \rightarrow y \in F$, since F is an associative filter w.r.t x for any $x \neq O$, then $x \in F$, so $y \in F$ and $y \rightarrow z \in F$ because F is a filter, then $z \in F, x \rightarrow z \in F$, so F is an implicative filter.

Remark 3.2. (1) If F is an associative filter w.r.t $x (x \in L)$, it may be not an implicative filter. For instance, $F = \{a, I\}$ is an associative filter w.r.t I instead of an implicative filter in Example 2.1.

(2) The converse of Theorem 3.3 is not always true. For instance, the filter $\{c, b, I\}$ is an implicative filter instead of an associative filter of \mathcal{L} in Example 2.1.

Theorem 3.4[1]. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, the following conclusions hold:

- (i) Every associative filter w.r.t x contains x itself.
- (ii) Every associative filter of \mathcal{L} is a filter.

Remark 3.3. If $\mathcal{L}=\{O, I\}$, $\{I\}$ is associative filter of \mathcal{L} . For any other lattice implication algebra \mathcal{L} , there doesn't exist associative filter of \mathcal{L} .

Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $|\mathcal{L}| \geq 2$. If F is an associative filter of \mathcal{L} , then $F = \{x \in L \mid x \neq O\}$. Hence $x, x' \in F$, it is a contradiction with F is a proper filter of \mathcal{L} .

Theorem 3.5. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $n \in N$, if F is a n -fold associative filter of \mathcal{L} , then F is a n -fold implicative filter of \mathcal{L} .

Proof. F is a n -fold associative filter of \mathcal{L} , For any $x \neq O$, $y, z \in L$, let $x^n \rightarrow (y \rightarrow z) \in F$, $x^n \rightarrow y \in F$ implies $z \in F$ for $z \leq x^n \rightarrow z$, it follows $x^n \rightarrow z \in F$. so F is a n -fold implicative filter of \mathcal{L} .

The relationship between n -fold associative filter and n -fold implicative filter is similar to the relationship between associative filter and implicative filter.

Theorem 3.6[5]. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. If F is an implicative filter and a prime filter, then F is an ultra-filter.

An implicative filter may not be a prime filter. In Example 2.3, we can verify that $\{I\}$ is an implicative filter instead of a prime filter. Because $a \vee b = I$, but $a \notin \{I\}$ and $b \notin \{I\}$.

A prime filter may not be an implicative filter. In Example 2.1, we can verify that $\{I, a\}$ is a prime filter instead of an implicative filter. Because $d \rightarrow (b \rightarrow c) = I$ and $d \rightarrow b = I$, but $d \rightarrow c = b \notin \{I, a\}$.

Theorem 3.7[5]. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, F is a proper filter of \mathcal{L} . If $x \in F$, then $x' \notin F$ for any $x \in L$.

Remark 3.4. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, then the following conclusions hold[1]:

- (1) The concept of implicative filter and positive implicative filter and Boolean filter are coinciding.
- (2) An ultra-filter is an obstinate filter. An Obstinate filter is a maximal proper filter.
- (3) Obstinate filter and maximal implicative filter are equivalent.

Theorem 3.8. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, then the maximal implicative filter of \mathcal{L} is the maximal proper filter.

Proof. Let F is the maximal implicative filter. So it is a Boolean filter, for any $x \in L, x \vee x' \in F$.

- (1) If $x \in F$ and $x' \in F$, then $F = L$.
- (2) If $x \in F$ and $x' \notin F$, then F is the maximal proper filter and F is a prime filter.

Theorem 3.9[5]. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, for any $a \in L, a \neq I$. If F is a submaximal filter with respect to a , then F is a prime filter.

Theorem 3.10. F is a prime filter, then there exist $a, a \in L$ and $a \notin F$, F is a submaximal filter with respect to a .

Proof. F is a prime filter, then F is an proper filter, there exist $a, a \in L$ and $a \notin F$, F is the maximal filter without a , it follows that F is a submaximal filter with respect to a .

Theorem 3.11. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, if it is a chain, then it has no dual atom or only one dual atom. And its trivial filters $\{I\}$ and L are involution filters.

Theorem 3.12. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, denote the set of dual atoms of L as $\text{Dat}(\mathcal{L})$, $F = \{x, I\}$ ($x \in \text{Dat}(\mathcal{L})$) is a filter, then F is an involution filter.

Proof. Let $F = \{x, I\}$ ($x \in \text{Dat}(\mathcal{L})$) be a filter. First, we prove that $F^* = L \setminus [O, x]$. In fact, for any $y \in L \setminus [O, x]$, if $y \vee x = x$, then $y \leq x$, which is a contradiction. It follows that $y \vee x = I$ and $y \in F^*$. Conversely, if $y \in F^*$, then $y \vee x = I$ and $y \in L \setminus [O, x]$.

Then we prove that $(F^*)^* = F$, $(F^*)^* = \{z \mid z \vee y = I, \text{ for any } y \in F^*\}$. Obviously, $F \subseteq (F^*)^*$, if there exist $b \in (F^*)^*$, and $b \notin F$, then $b \in [O, x]$ or $b \in L \setminus \{[O, x], I\}$, $b \vee y = I$ for any $y \in F^*$. It does not exist $b \in [O, x]$ with $b \vee y = I$ for any $y \in F^*$. If $b \in L \setminus \{[O, x], I\}$, exist $c \in F^*$, and $c \leq b$, but $b \vee c \neq I$, which is a contradiction. So $(F^*)^* = F$.

Example 3.1. As for the lattice implication algebra defined in example 2.1, it is routine to check that $\{a, I\}$ is a filter, and it is an involution filter.

Let \mathcal{L} be a lattice H implication algebra, for any $x \in L$, $x \wedge x' = O, x \vee x' = I$, then any filter of \mathcal{L} is an involution filter.

Theorem 3.14[1]. If F is a prime filter of lattice H implication algebra \mathcal{L} , then F is an ultra filter.

Theorem 3.15. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice H implication algebra, then all kinds of filter in \mathcal{L} are equivalent.

Remark 3.5

- AF associative filter w.r.t x , for any $x \in L$
 - OF obstinate filter
 - BF Boolean filter
 - IF implicative filter
 - PF prime filter
 - UF ultra-filter
 - n-AF n-fold associative filter
 - maxIF maximal implicative filter
 - maxTF maximal proper filter
 - submaxF submaximal filter w.r.t x
 - PIF positive implicative filter
 - InF involution filter
 - n-IF n-fold implicative filter
- implicative filter \Leftrightarrow Boolean filter \Leftrightarrow positive implicative filter

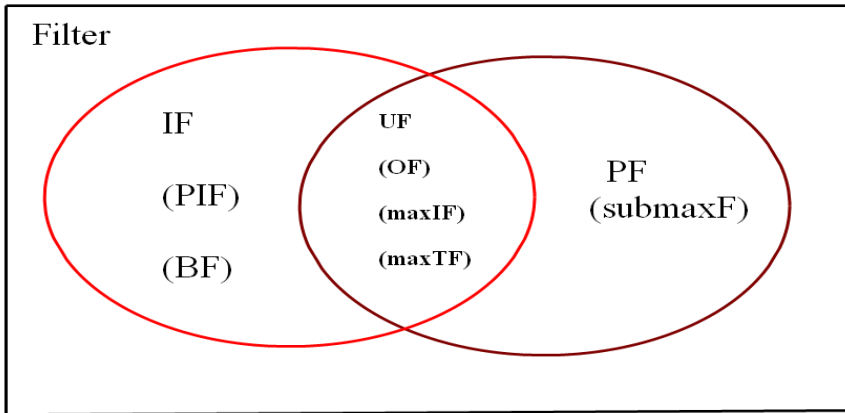


Fig. 4. The relationship of filters in lattice implication algebra

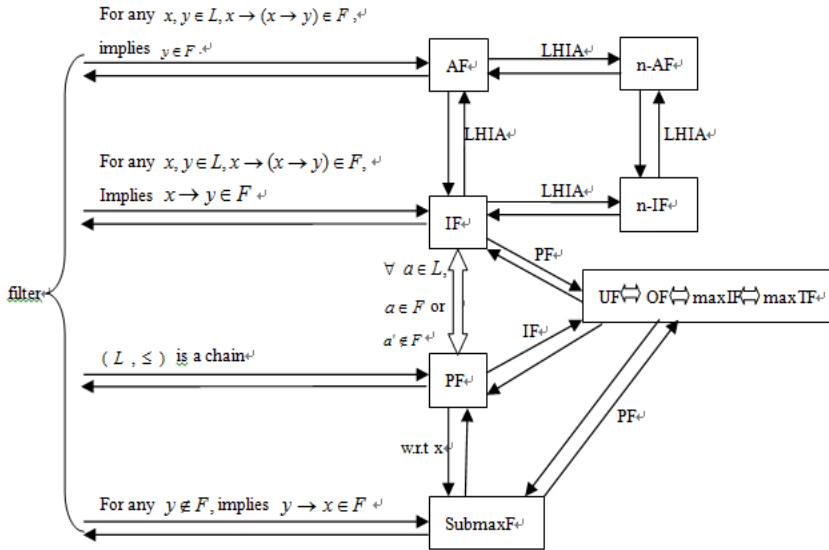


Fig. 5. Transformation conditions among these filters in lattice implication algebra

4 Filter Lattice

Let $\Sigma(L)$ the set of all the filters in LIA, define the partial order and binary operator \wedge, \vee as follows: for any $F_1, F_2 \in \Sigma(L)$, $F_1 \leq F_2$ if and only if $F_1 \subseteq F_2$, $F_1 \wedge F_2 = F_1 \cap F_2$, $F_1 \vee F_2 = [F_1 \cup F_2]$, then $(\Sigma(L), \leq)$ is a lattice, L and $\{I\}$ are the greatest and the smallest elements of $\Sigma(L)$.

Theorem 4.1. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a finite lattice implication algebra, then \mathcal{L} can be respect with the product of finite chain $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_n$, I_i is the largest element of L_i ($0 \leq i \leq n$), the there are 2^n filters in L , and
$$\sum(L) = \left\{ \prod_{i=1}^n a_i L_i \mid a_i L_i = \begin{cases} \{I_i\}, a_i = O_i, \\ L_i, a_i = I_i \end{cases}, i = 1, 2, \dots, n \right\}.$$

Theorem 4.2. Let $\mathcal{L}=(L, \vee, \wedge, ', \rightarrow, O, I)$ be a finite lattice implication algebra, then the filter lattice $(\sum(L), \leq)$ of \mathcal{L} is isomorphic to Boolean lattice.

Proof. Let \mathcal{L} be a finite lattice implication algebra, \mathcal{L} can be respect with the direct product of finite chain, let $\mathcal{L}=\mathcal{L}_1 \times \mathcal{L}_2 \times \dots \times \mathcal{L}_n$, \mathcal{L}_i ($0 \leq i \leq n$) is finite chain, for any \mathcal{L}_i , there is an isomorphic mapping f_i ($0 \leq i \leq n$) form $\sum(L_i)$ to B_2 , so the mapping $f : \sum(L) \rightarrow B_2^n$

$$\begin{aligned} f(\sum(L)) &= f(\sum(L_1) \times \sum(L_2) \times \dots \times \sum(L_n)) \\ &= f_1(\sum(L_1)) \times f_2(\sum(L_2)) \times \dots \times f_n(\sum(L_n)) \\ &= B_2 \times B_2 \times \dots \times B_2 \\ &= B_2^n \end{aligned}$$

So $(\sum(L), \leq)$ of \mathcal{L} is isomorphic to Boolean lattice.

Acknowledgments. This work is supported by the National Natural Science Foundation of P.R. China (Grant No. 60875034).

References

1. Xu, Y., Ruan, D., Qin, K.Y., Liu, J.: Lattice-Valued Logic—An Alternative Approach to Treat Fuzziness and Incomparability. Springer, New York (2003)
2. Song, S.Z., Jun, Y.B.: On n-fold Positive Implicative Filters Of Lattice Implication Algebras. Soochow J. Math. 30(2), 225–235 (2004)
3. Van Gasse, B., Deschrijver, G., Cornelis, C., Kerre, E.E.: Filters of residuated lattices and triangle algebras. Information Sciences 180, 3006–3020 (2010)
4. Wu, M.H.: Study on the lattice implication algebras and Correlative Substructures. Master Thesis, Southwest Jiaotong University (2010)
5. Du, S.K.: On lattice implication algebra and the relation between it and correlative logic algebras. Master thesis, Southwest Jiaotong University (2010)
6. Xu, Y., Qin, K.Y.: On filters of lattice implication algebras. J. Fuzzy Math. 1, 251–260 (1993) (in Chinese)
7. Jun, Y.B., Xu, Y., Qin, K.: Positive implicative and associative filters of lattice implication algebras. Bull. Korean Math. Soc. 35(1), 53–61 (1998)

8. Meng, B.: Prime filters of lattice implication algebra. *The Journal of Northwest University* 28(3), 189–192 (1998) (in Chinese)
9. Qin, K., Xu, Y.: On ultra-filters of lattice implication algebra. *The Journal of Southwest Jiaotong University* 34(1), 189–192 (1999) (in Chinese)
10. Wang, X., Xu, Y., Song, Z.: Some properties of filters in lattice implication algebra. *The Journal of Southwest Jiaotong University* 36(5), 536–539 (2001) (in Chinese)
11. Xu, Y., Ruan, D., Qin, K., Liu, J.: *Lattice-valued Logic*. Springer, Heidelberg (2003)
12. Song, S.Z., Jun, Y.B.: On n -fold positive implicative filters of lattice implication algebras. *Soochow Journal of Mathematics* 30(2), 225–235 (2004)

Effects of Vision Field on Evolution of Cooperation among Mobile Agents

Wei-Ye Wang, Zhou Zhou, and Xiao-Long Jiang

School of Electronic and Information Engineering, Beihang University,
Beijing, 100083, P.R. China
afatfox@163.com

Abstract. We introduce the heterogeneous vision radius which follows power-law distribution into the evolutionary prisoner's dilemma game among mobile agents to study the evolution of cooperation. It is found that the cooperation level of the system presents remarkable differences when vision radius follows power-law distribution with different exponents. The cooperative agents can survive in such a system provided that the heterogeneity of vision radius is moderate and the temptation of defection is not too high. This work may be beneficial to understand cooperative behavior in biological and social systems consisted of mobile agents.

Keywords: Prisoner's dilemma game, Cooperation, Vision radius, Mobility.

1 Introduction

Cooperation is ubiquitous in biological system and human society. Darwin's theory of evolution considers that there exists brutal competition among selfish individuals so that natural selection doesn't favor cooperators. So far, in biology and physics, much investigation has been focused on the problem how cooperation emerges and sustains among selfish individuals. The evolutionary game theory is considered as a powerful theoretical framework to study the cooperative behavior[1-4]. Especially, the prisoner's dilemma game (PDG) together with evolutionary context is widely investigated by researchers from different communities (e.g., mathematics, automation, computer science, physics, biology, etc.) who want to reveal the origin of altruistic behaviors emerging in selfish populations[5-9].

The classical PDG is a two-player game where each player can adopt one of two optional strategies: cooperation (C) and defection (D). For mutual cooperation (defection), each player obtains R (P). If one is cooperator and the other chooses defection, the defector can get the highest temptation T and the cooperator can only get the sucker's payoff S. The above four elements must satisfy two strict rules: $T > R > P > S$ and $2R > T + S$. One can find that these rules inevitably result in a cooperation dilemma: for a one-shot game, the better choice is always defect no matter what the opponent's selection is, but the outcome of mutual cooperation (defection) is the highest (lowest). If the agents are well mixed, the system will soon fall into the pure-defector state.

The evolutionary games have attracted much attention since Nowak and May reported surprising cooperation emergence and amazing spatial chaos could be induced

by the evolutionary PDG in a two-dimensional lattice[10]. After that, more effort has been devoted to investigating the different mechanisms which sustain cooperation in spatial structure, including regular networks[11, 12], small-world networks[13, 14], scale-free networks[15-18] and adaptive networks[19-21]. In addition, many works have been also done to explain the cooperative behavior in following aspects: memory effects[22, 23], reward and punishment mechanism[24-26], environment or payoff noise[27-29], the preferential selection[30], time scale effects[31], aspiration-based or asymmetric payoff effects[32-34] and system heterogeneity effects [35-37], etc. For more details about evolution of cooperation, one can refer to two comprehensive reviews [38].

Mobility is a fundamental characteristic of individuals in the real world, and it plays a crucial role in various dynamical processes[39, 40]. Thus researches on evolutionary games are naturally extended to study the evolution of cooperation among mobile agents. In the scene of Meloni et.al., agents randomly move in the two-dimensional plane and play games with agents within their vision field. They found cooperation can survive if the temptation to defect and the velocity of agents are not too high[41]. Yang et.al. proposed an aspiration-induced migration model where individuals can decide whether to move according to the aspiration level and it is found that the cooperative behavior will be best favored by the moderate aspiration level[42, 43]. Helbing et.al. proposed a success-driven migration model, where a surprising sudden outbreak of cooperation is induced in the noisy system[44].

However, in most previous works, the vision field is the same for all agents, which somewhat contradicts many real observations, e.g., a myopic person cannot see so far as a man with good eyesight and a soaring eagle surely has a much larger vision field than a little sparrow. Thus in this paper, we investigate the effect of heterogeneous view field via comparing different exponents of vision radius following power-law distributions on the cooperative evolution. It is found that different power-law distribution can induce discrepant cooperative behaviors.

Our paper is organized as follows. The model is described in Section 2. The simulation results and discussions are present in Section 3. Finally, we conclude the work in Section 4.

2 The Model

Our model is considered to include N individuals moving in a two-dimensional square plane of size $L \times L$ with periodic boundary restrictions and playing prisoner's dilemma game on the instantaneous networks of contacts. Initially, all agents are randomly located in the plane and their directions of motion are distributed with uniform probability in $[-\pi, \pi]$.

First we will present the moving rule of individuals. All agent move at the same absolute velocity v during the whole evolution, but can change the motion direction $\theta_i(t)$ ($i=1,2,\dots,N$). We define $\vec{x}_i(t)$ as the position of the agent i at time step t , $t=0,1,2,\dots$, and updated according to following expression at each time step:

$$\vec{x}_i(t+1) = \vec{x}_i(t) + v e^{i\theta_i(t)}. \quad (1)$$

Next we will introduce the process of evolutionary game. We can simplify the payoff matrix of PDG as follows [6]: $R=1$, $P=S=0$ and $T=b$, where $1 \leq b \leq 2$ is defined as the temptation to defection. Initially, all agents are randomly assigned one of the two available strategies which is cooperation (C) or defection (D) with equal probability. In each time step, agent i ($i=1,2,\dots,N$) plays PDG with all its neighbors (here neighbors are defined as agents in i 's vision field), and accumulates payoffs as the total payoff of agent i at each round. Then, agent i imitates the strategy of its richest neighbor as i 's strategy in the next time step if the payoff of agent i is less than that of the richest one. Otherwise, agent i holds its own strategy unchanged in the next time step. At time step t , the vision field of agent i is expressed as $R_i(t)$:

$$R_i(t) = \left\{ \left\| \vec{x}_i(t) - \vec{x}_j(t) \right\| \leq r_i, j \in N, j \neq i \right\}, \quad (2)$$

i.e. a circle of view radius r_i ($i=1,2,\dots,N$) centered on agent i , where $\left\| \vec{x}_i(t) - \vec{x}_j(t) \right\|$ is the Euclidean distance between i and j in the $L \times L$ plane. In the case of original model, r_i is the same for all agents. But in real world, the range that individuals interact with others is usually heterogeneous. Thus in our model, view radius r_i is considered to be different from each other and it is rescaled as:

$$r_i = x^{\frac{1}{n}} - \frac{n}{n-1} + \langle r \rangle, n > 1, \quad (3)$$

where vision radius r_i follows power-law distribution with exponent n and x is random number of uniform distribution in the interval $[0, 1]$. It is noteworthy that $\int_0^1 r_i(x) dx = \langle r \rangle$ guarantees the average value of all individuals' view radius in population is $\langle r \rangle$.

3 Results and Discussion

All the simulations below are carried out on the 40×40 square plane with periodic boundary conditions. The number of agents in population is set as $N = 1000$, and the average view radius of whole population is set as $\langle r \rangle = 3$. The equilibrium fraction of cooperators is obtained by averaging the last 10,000 time steps after the whole evolution process of 100,000 generations. All results are averaged over 50 simulations with individual realizations.

Fig. 1(a) shows the relationship between the frequency of cooperators f_c and temptation of defection b . Generally, the level of cooperator monotonically decreases with the increase of b . Compared to that of the original prisoner's dilemma game, the cooperation level is highly restrained in the case of $n=1.5$ but remarkably promoted when n is larger (e.g., $n=4, 6, 8$). The relationship between the frequency of cooperators f_c and the exponent n of power-law distribution is plotted in Fig 1(b). For a fixed temptation of defection b , it is obvious that the relationship of n and the cooperation levels is non-monotonic. The defectors are dominant in population which the level of

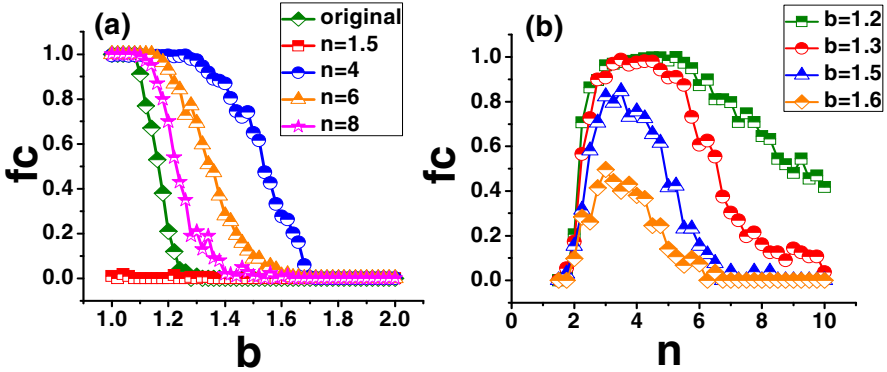


Fig. 1. (a) The relationship between the frequency of cooperators fc and temptation of defection b . (b) The relationship between the frequency of cooperators fc and the exponent n of power-law distribution. Here, $N=1000$, $L=40$, $v=0.2$, $\langle r \rangle=3$.

cooperation fc is less than 10% if $n \leq 2$. With the increment of n , the density of cooperators is sharply promoted until n reaches a specific threshold ($n \approx 4$). After that, the cooperation level quickly falls.

Many previous works have revealed that the degree distribution plays a crucial role in the evolution of cooperation [14, 15, 45, 46]. In the mobile model, the degree of agents is closely related to the vision radius. Since the change of n can generate the diversity of view radius, we attribute the promotion of cooperation to the heterogeneous degree distributions induced by the heterogeneous view radius. First, we explore the distributions of view radius under different n (Fig 2). Obviously, with the increment of n , the heterogeneity of view radius among individuals is weakened. Fig 3 plots the relation between individual's degree and view radius. It is found that the degree distributions are well in accordance with the distribution of the view radius.

When $n=1.5$ (Fig 3(a)), the correlation of degree and view radius follows a power-law relationship and the heterogeneity of degree distribution is the largest (the range of degree from 1 to 999). If $n=4$ (Fig 3(b)), although degree and view radius is still power-law correlated, the heterogeneity of degree distribution is weaker (the range of degree from 1 to 310). When $n=8$ (Fig 3(c)), the relation between the degree and view radius has not been power-law type and the degree distribution is a little more heterogeneous than that of the original version (Fig 3(d)) where the view radius is constant.

Fig 4 shows agents' degree distributions under different values of n . In the case of $n=1.5$ (Fig 4(a)), the degree distribution approaches to the power-law distribution and takes on high heterogeneity. When $n=4$ (Fig 4(b)), the degree distribution owns a power-law tail (degree > 20). As n increases, degree distributions gradually tend to normal distribution. (Fig 4(c)(d)). Namely, the increase of n reduces heterogeneity of view radius, resulting in more homogeneous degree distribution.

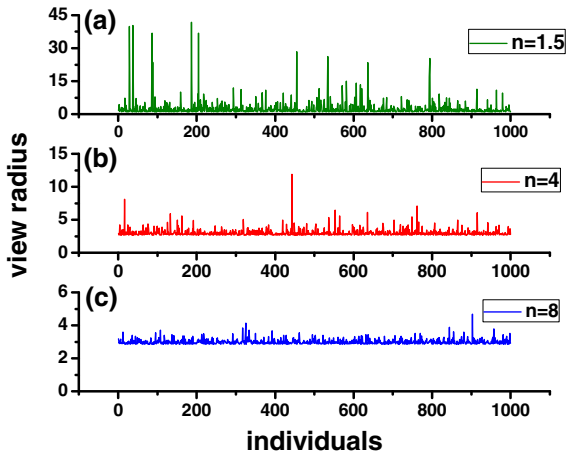


Fig. 2. The distributions of view radius r_i form=1.5, 4 and 8. Here $\langle r \rangle = 3$.

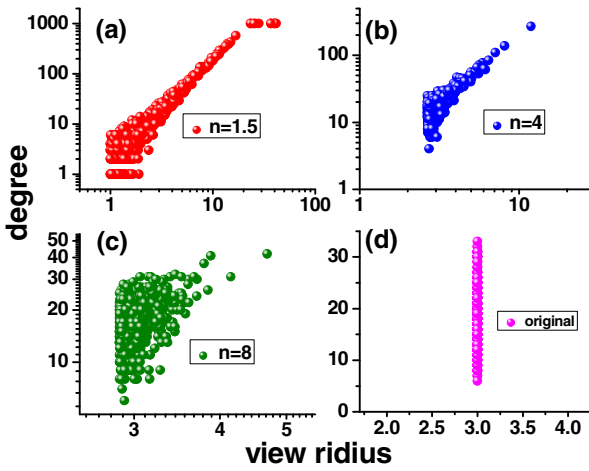


Fig. 3. The relation between agent's degree and view radius. Here, $N=1000$, $L=40$, $v=0.2$, $\langle r \rangle = 3$.

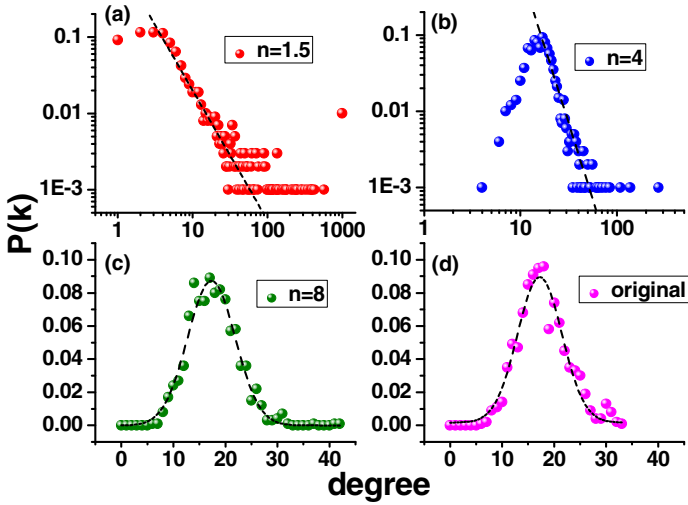


Fig. 4. The degree distributions under different values of n . Here, $N=1000$, $L=40$, $v=0.2$, $\langle r \rangle=3$, $b=1.3$. It is noteworthy that (a) and (b) are based on log-log coordinate system, and (c) and (d) are based on linear coordinate system.

Next, we will discuss the effect of n on cooperative behaviors among mobile agents in detail. Previous works, have confirmed that the cooperative behavior can be enhanced by the hub nodes in scale-free networks (the hub nodes can influence their neighbors into cooperators to guarantee their long-time success)[47]. In our mobile model, the hub nodes also have the capability in forming cooperator clusters but the mobile cooperator cluster is not so stable as the static cooperator clusters. In the static case, the compact cooperator cluster can get high payoffs to resist the invasion of defectors; whereas in the mobile case, a single cooperator cluster can be easily diluted during the unordered movement and the cooperators can hardly sustain. When n is small ($n=1.5$), the degree distribution is highly heterogeneous and the number of hub node is quite small. In the initial state (Fig 5(a)), cooperators and defectors are randomly mixed (red represents cooperators and blue represents defectors). After 50 time steps, some small cooperator clusters emerged (Fig 5(b)). Due to agents' mobility, the cooperator clusters gradually disappear (Fig 5(c)(d)). In a word, highly heterogeneous degree distribution and agents' motion lead to difficult sustainment of cooperative behavior. When the value of n is moderate ($n=4$), the heterogeneity of players' degree distribution is not too high compared to that of $n=1.5$, and the number of hub nodes is much more than that of $n=1.5$. As shown in Fig 5(e)-(h), numerous small cooperator clusters can merge to form huge cooperator clusters, and the huge cooperator clusters can better resist the invasion of defectors even under the mobile case (it is so huge

that most cooperators can only move inside it to get high payoffs). As the value of n increases ($n=8$), the degree distribution gradually tends to normal distribution and becomes more and more homogeneous so that hub nodes, on which the cooperator clusters rely, gradually disappear. Consequently, it is not surprising to have a low cooperation level for large values of n (Fig 5(i)-(l)).

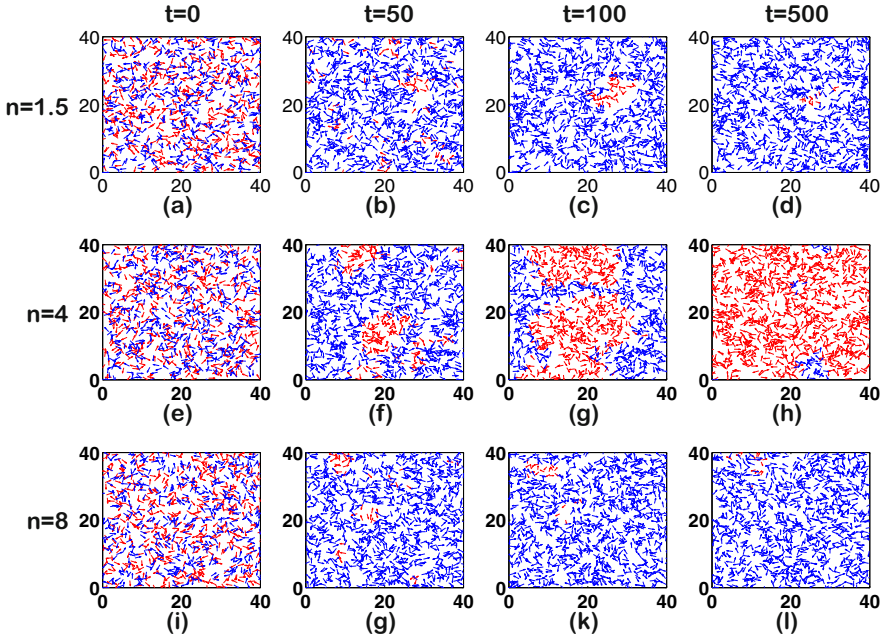


Fig. 5. The spatial state of cooperative evolution process for different n at time step $t=0, 50, 100$ and 500 . Here, $N=1000, L=40, v=0.2, \langle r \rangle=3, b=1.3$. Red represents cooperators and blue represents defectors.

To further confirm the role of hub nodes in the evolution of cooperation, we select top ten agents with the largest view radius to fix their strategies as C (expressed by square green line in Fig (6)) or D (expressed by round red line in Fig (6)) during the whole evolution and investigate the cooperative behavior of the system (for the original version, we stochastically pick 10 nodes to fix their strategies as C or D). The results are shown in Fig 6. In the case of $n=1.5$ (Fig 6(a)), the system's cooperation level is remarkably promoted if the 10 agents with the largest vision radius are fixed as cooperators, e.g., when $b=1.0$, the cooperation level is promoted from 0.03 to 0.87. When $n=4$ (Fig 6(b)) and 8 (Fig 6(c)), the cooperative behaviors are also promoted or depressed, depending on the strategy of the 10 selected agents. In the original version (Fig 6(d)), since there is no hub node in the system, the cooperation level has no remarkable difference when 10 randomly selected agents are fixed as cooperators or defectors.

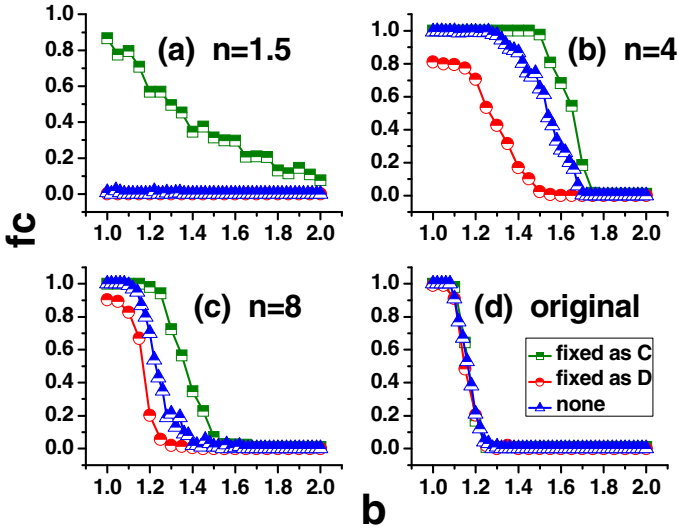


Fig. 6. We select ten hub nodes to fix their strategies into C or D during the whole evolution and research the cooperative behavior of the system. It is noteworthy that, in the figure (d), we randomly pick 10 nodes to fix their strategies as a result of the same view radius. Here, $N=1000$, $L=40$, $v=0.2$, $\langle r \rangle = 3$.

4 Conclusion

In summary, we have investigated the cooperative behaviors among mobile agents, where agents move randomly on a two-dimensional plane and play PDG with agents within their vision field. The vision radius of agents follows power-law distribution with different exponent n . We find that the highest cooperative level is induced by intermediate value of n . Then we give a comprehensive explanation for this interesting phenomenon via investigating the relationship of vision radius and degree, and the underlying mechanism of the extinction/emergence of mobile cooperator clusters. Our work may shed some light on the role of mobility in evolutionary game theory.

Acknowledgments. We thank Wen-Bo Du and Han-Xin Yang for useful discussions. This work is supported by the National Basic Research Program of China (Grant No. 2010CB731805), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (Grant No. 60921001), and the National Key Technologies R&D Program of China (Grant No. 2011BAH24B02).

References

- [1] Maynard Smith, J.: Evolution and the Theory of Games. Cambridge University Press, Cambridge (1982)
- [2] Axelrod, R.: The Evolution of Cooperation. Basic Books, New York (1984)

- [3] Hofbauer, J., Sigmund, K.: *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge (1998)
- [4] Nowak, M.A.: *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, Cambridge (2006)
- [5] Srivastava, V., Neel, J., Mackenzie, A.B., Menon, R., DaSilva, L.A., Hicks, J.E., Reed, J.H., Gille, R.P.: Using Game Theory to Analyze Wireless Ad Hoc Networks. *IEEE Commun. Surv. Tutorial* 7(4), 46–56 (2005)
- [6] Srinivasan, V., Nuggehalli, P., Chiasserini, C.F., Rao, R.R.: Cooperation in Wireless Ad Hoc Networks. In: *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications*, vol. 2, pp. 808–817 (2003)
- [7] Tomlin, C.J., Lygeros, J., Shankar Sastry, S.: A Game Theoretic Approach to Controller Design for Hybrid Systems. *Proceedings of the IEEE* 88(7), 949–970 (2000)
- [8] Hill, J.C., Johnson, F.R., Archibald, J.K., Frost, R.L., Stirling, W.C.: A Cooperative Multi-Agent Approach to Free Flight. In: *Proceedings of the 4th Int. Joint Conference on Autonomous Agents and Multi-agent Systems*, The Netherlands, pp. 1083–1090. ACM (2005)
- [9] Archibald, J.K., Hill, J.C., Jepsen, N.A., Stirling, W.C., Frost, R.L.: A Satisficing Approach to Aircraft Conflict Resolution. *IEEE Transactions Systems, Man, and Cybernetics, Part C: Applications and Reviews* 38(4), 510–521 (2008)
- [10] Nowak, M.A., May, R.M.: Evolutionary Games and Spatial Chaos. *Nature (London)* 359, 826–829 (1992)
- [11] Szabó, G., Tóke, C.: Evolutionary Prisoner’s Dilemma Game on a Square Lattice. *Phys. Rev. E* 58, 69–73 (1998)
- [12] Szolnoki, A., Perc, M.: Impact of Critical Mass on the Evolution of Cooperation in Spatial Public Goods Games. *Phys. Rev. E* 81, 057101 (2010)
- [13] Du, W.-B., Cao, X.-B., Yang, H.-X., Hu, M.-B.: Evolutionary Prisoner’s Dilemma on Newman-Watts Social Networks with an Asymmetric Payoff Distribution Mechanism. *Chin. Phys. B* 19, 10204 (2010)
- [14] Du, W.-B., Cao, X.-B., Zhao, L., Zhou, H.: Evolutionary Games on Weighted Newman-Watts Small-World Networks. *Chinese Physics Letters* 26, 58701 (2009)
- [15] Santos, F.C., Pacheco, J.M.: Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation. *Phys. Rev. Lett.* 95, 098104 (2005)
- [16] Cao, X.-B., Du, W.-B., Rong, Z.-H.: The Evolutionary Public Goods Game on Scale-Free Networks with Heterogeneous Investment. *Physica A* 389, 1273–1280 (2010)
- [17] Du, W.-B., Cao, X.-B., Hu, M.-B., Yang, H.-X., Zhou, H.: Effects of Expectation and Noise on Evolutionary Games. *Physica A* 388, 2215–2220 (2009)
- [18] Du, W.-B., Zheng, H.-R., Hu, M.-B.: Evolutionary Prisoner’s Dilemma Game on Weighted Scale-Free Networks. *Physica A* 387, 3796–3800 (2008)
- [19] Zimmermann, M.G., Eguíluz, V.M., Miguel, M.S.: Coevolution of Dynamical States and Interactions in Dynamic Networks. *Phys. Rev. E* 69, 065102 (2004)
- [20] Fu, F., Hauert, C., Nowak, M.A., Wang, L.: Reputation-Based Partner Choice Promotes Cooperation in Social Networks. *Phys. Rev. E* 78, 026117 (2008)
- [21] Poncela, J., Gómez-Gardeñes, J., Floría, L.M., Sánchez, A., Moreno, Y.: Complex Cooperative Networks from Evolutionary Preferential Attachment. *PLoS One* 3(6), e2449 (2008)
- [22] Wang, W.-X., Ren, J., Chen, G.-R., Wang, B.-H.: Memory-Based Snowdrift Game on Networks. *Phys. Rev. E* 74, 056113 (2006)
- [23] Du, W.-B., Cao, X.-B., Liu, R.-R., Jia, C.-X.: The Effect of a History-Fitness-Based Updating Rule on Evolutionary Games. *International Journal of Modern Physics C* 21, 1433–1442 (2010)

- [24] Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A., Sigmund, K.: Via Freedom to Coercion: The Emergence of Costly Punishment. *Science* 316, 1905–1907 (2007)
- [25] Ohtsuki, H., Iwasa, Y., Nowak, M.A.: Indirect Reciprocity Provides Only a Narrow Margin of Efficiency for Costly Punishment. *Nature* 457, 79–82 (2009)
- [26] Wang, Z., Xu, Z.-J., Zhang, L.-Z.: Punishment in Optional Public Goods Games. *Chin. Phys. B* 19, 110201 (2010)
- [27] Vukov, J., Szabó, G., Szolnoki, A.: Cooperation in the Noisy Case: Prisoner’s Dilemma Game on Two Types of Regular Random Graph. *Phys. Rev. E* 73, 67103 (2006)
- [28] Szolnoki, A., Vukov, J., Szabó, G.: Selection of Noise Level in Strategy Adoption for Spatial Social Dilemmas. *Phys. Rev. E* 80, 56112 (2009)
- [29] Szabó, G., Szolnoki, A., Vukov, J.: Selection of Dynamical Rules in Spatial Prisoner’s Dilemma Games. *Europhys. Lett.* 87, 18007 (2009)
- [30] Du, W.-B., Cao, X.-B., Zhao, L., Hu, M.-B.: Evolutionary Games on Scale-Free Networks with a Preferential Selection Mechanism. *Physica A* 388, 4509–4514 (2009)
- [31] Szolnoki, A., Perc, M., Szabó, G., Stark, H.-U.: Impact of Aging on the Evolution of Cooperation in the Spatial Prisoner’s Dilemma Game. *Phys. Rev. E* 80, 021901 (2009)
- [32] Zhang, J., Fang, Y.-P., Du, W.-B., Cao, X.-B.: Promotion of Cooperation in Aspiration-Based Spatial Prisoner’s Dilemma Game. *Physica A* 390, 2258–2266 (2011)
- [33] Du, W.-B., Cao, X.-B., Hu, M.-B., Wang, W.-X.: Asymmetric Cost in Snowdrift Game on Scale-Free Networks. *Europhys. Lett.* 87, 60004 (2009)
- [34] Du, W.-B., Cao, X.-B., Hu, M.-B.: The Effect of Asymmetric Payoff Mechanism on Evolutionary Networked Prisoner’s Dilemma Game. *Physica A* 388, 5005–5012 (2009)
- [35] Zhang, J., Wang, W.-Y., Du, W.-B., Cao, X.-B.: Evolution of Cooperation among Mobile Agents with Heterogenous View Radii. *Physica A* 390, 2251–2257 (2011)
- [36] Du, W.-B., Zhou, H., Liu, Z., Cao, X.-B.: The Effect of Pinning Control on Evolutionary Prisoner’s Dilemma Game. *Modern Physics Letters B* 24, 2581–2589 (2010)
- [37] Du, W.-B., Cao, X.-B., Zheng, H.-R., Zhou, H., Hu, M.-B.: Evolutionary Games in Multi-Agent Systems of Weighted Social Networks. *International Journal of Modern Physics C* 20, 701–710 (2009)
- [38] Szabó, G., Fáth, G.: Evolutionary Games on Graphs. *Phys. Rep.* 446, 97–216 (2007)
- [39] Hufnagel, L., Brockmann, D., Geisel, T.: The Scaling Laws of Human Travel. *Nature* 439, 462–465 (2006)
- [40] González, M.C., Hidalgo, C.A., Barabási, A.L.: Understanding Individual Human Mobility Patterns. *Nature* 453, 779–782 (2008)
- [41] Meloni, S., Buscarino, A., Fortuna, L., Frasca, M., Gómez-Gardeñes, J., Latora, V., Moreno, Y.: Effects of Mobility in a Population of Prisoner’s Dilemma Players. *Phys. Rev. E* 79, 067101 (2009)
- [42] Lin, Y.-T., Yang, H.-X., Wu, Z.-X., Wang, B.-H.: Promotion of Cooperation by Aspiration-Induced Migration. *Physica A* 390, 77–82 (2011)
- [43] Yang, H.-X., Wu, Z.-X., Wang, B.-H.: Role of Aspiration-Induced Migration in Cooperation. *Phys. Rev. E* 81, 065101(R) (2010)
- [44] Helbing, D., Yu, W.: The Outbreak of Cooperation among Success-Driven Individuals under Noisy Conditions. *Proc. Natl. Acad. Sci. USA* 106, 3680–3685 (2009)
- [45] Gómez-Gardeñes, J., Campillo, M., Floría, L.M., Moreno, Y.: Dynamical Organization of Cooperation in Complex Topologies. *Phys. Rev. Lett.* 98, 108103 (2007)
- [46] Poncela, J., Gómez-Gardeñes, J., Floría, L.M., Moreno, Y.: Robustness of Cooperation in the Evolutionary Prisoner’s Dilemma on Complex Networks. *New J. Phys.* 9, 184 (2007)
- [47] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U.: Complex Networks: Structure and Dynamics. *Phys. Rep.* 424, 175–308 (2006)

Deviating from Common Context in Individual Semiosis in Multi-Agent Systems

Wojciech Lorkiewicz^{1,2}, Radoslaw Katarzyniak¹, and Ryszard Kowalczyk²

¹ Wroclaw University of Technology, Institute of Informatics
{wojciech.lorkiewicz,radoslaw.katarzyniak}@pwr.wroc.pl

² Swinburne University of Technology, FICT
{wlorckiewicz,rkowalczyk}@groupwise.swin.edu.au

Abstract. In order to communicate successfully highly distributed agents must utilise a shared naming convention. Such a naming convention can be developed by the agents through the process of semiosis where the agents collectively establish the naming convention for the objects. Narrowing the interaction pattern of Language Game Model to a single speaker (teacher) we model the process of individual semiosis. Further, using a developed simulation framework of the process, we analyse the dynamic behaviour of the alignment against deviations from the idealised settings of shared context. In particular, we study the reaction on the formation of naming convention among the interacting population in settings where the communication channels are subject to imperfection. As such, this research fills the current gap and investigates the dynamics of the fine-grained model of semiosis. In particular, we analyse both, analytically and through simulation, the observed phase transition in the alignment process.

Keywords: multi-agent systems, individual semiosis, language game.

1 Introduction

The fundamental layer of language requires manipulation of grounded symbols [5], allowing an individual to relate experienced sensory patterns to an arbitrary established system of signs - language [8]. The meaning of the utilised signs goes beyond an individual and all of the internal linguistic structures are additionally influenced by the collective. Consequently, the agents must not only be able to shape and develop their individual linguistic structures but also be able to align them with other agents. Moreover, the heterogeneity and autonomy of the system implies that the decisions are highly distributed and symbols are shaped individually by the agents. As such a group of individuals engaged in communication activity can be considered as a dynamic distributed cognitive system' [9], where different participants actively align their linguistic systems in order to collaborate, i.e. perform a certain distributed, social or cognitive task. Developing an arbitrary language in a highly distributed and dynamic population of autonomous agents is not a trivial task [6]. In the literature there already

exist several approaches to this problem. The most relevant research involves the Language Game Model (LGM) [9,10] oriented towards the development of a general and consistent language framework. Steels et al. [11] have introduced an intuitive and common-sense interaction pattern, i.e., a routinised game that is played by two agents, that defines the fundamental interplay within the system. Despite its popularity, only a limited set of models of the LGM have been investigated, i.e., the main focus has been on studies involving the idealised settings with significant feedback and limited context sizes. Several highly rational scenarios have been neglected, in particular rational ones in terms of embodied application, e.g. robotics and smart sensor networks. Presented research tries to fill this gap by studying the dynamics of the process of individual semiosis that utilises the fine-grained model of language acquisition in multi-agent setting. Assuming basic capabilities of the agents (Sec.3), i.e., autonomous object-oriented perception, ability to utter words and ability to register words, we modify the original interaction pattern of LGM to a single and fixed speaker (teacher) in order to deal with the model of individual semiosis [6]. In particular, we study the behaviour of the alignment process, in the settings where the channels are subject to imperfection and deviate from the shared context restriction.

This paper addresses a sub-problem of language alignment in a population of multiple, distributed and autonomous individuals (Sec.2). It introduces the internal organisation of the system, i.e. agent, interaction and alignment strategies, and the proposed measures of system's performance (Sec.3). Further, it presents an analysis of the simulated dynamic behaviour of this model (Sec.4) in the settings of sporadic deviation from the shared context settings (Sec.5).

2 Individual Semiosis and the Naming Game

Individual semiosis is a sub-process of the process of semiosis [6], i.e. a group of processes that are involved in the task where a population of autonomous agents establishes a common language of communication. In the individual semiosis setting a learning agent is shaping (aligning) its language with a distinguished external source of meaning, i.e., agents with a predefined and static meaning of utilised language symbols – static language¹. Through numerous learning episodes individuals are able to fine-grain the correct relation between words and objects [9,2,3], in particular due to the multi-modal nature of interaction between learner agent and a teacher agent. Through the co-occurrence of a particular sensor activation pattern and a particular linguistic sign the learning agent is able to develop the proper correlations between a word and an external object [5]. Such a distributed mechanism of alignment allows the entire population to identify the consistent source of meaning for each of the utilised symbols and results in the spread of its successful usage, even for large lexicons [2].

¹ Following a basic simplification that word learning is, in principle, a rather simple task of mapping linguistic labels (words) onto a set of pre-established concepts (objects) [1] we further assume that the term language relates to a naming convention represented by a certain lexicon.

Individual semiosis represents a narrower process that focuses on each learner independently. As aforementioned, through numerous interactions the learning agent is trying to identify the correct, i.e., consistent with the teacher, sources of meaning for each of the encountered language symbols. In other words, the learning agent (learning population) needs to align its individual naming convention with the naming convention incorporated by the established teacher (mature population). Moreover, due to the existence of the mature population (population of teachers with predefined naming conventions) the alignment depends solely on the interaction between the agents of two types. This intuitive fact is significant, as it allows to idealise the process by neglecting the internal interactions within the learning (and mature) population and focus solely on the communication between learners and teachers.

In the LGM a single interaction involves an interplay between a random pair of agents, one acting as a speaker, whilst the other as a hearer. Both interacting agents are presented with a scene comprised of a set of selected objects – context – and both perceive the same context. Then the speaker randomly selects an object from its sight and utters its name, according to its individual linguistic mechanism. Whilst the hearer, uses the received utterance as a clue to identify the intended object in the set of available objects in the current context. Further, three basic types of the game can be identified [14]. In the simplest case, both agents receive feedback – the hearer points to the intended interpretation and the speaker points to the intended topic. In the case of limited feedback only the speaker receives feedback – the hearer points to the intended interpretation. In the no-feedback case neither the speaker nor the hearer receives additional feedback after the game – leaving both agents clueless about the results of their interaction.

Interaction pattern of LGM assumes that agents are equipped with a predeveloped ability to establish shared context (joint attention) and a pointing mechanism (feedback). Maintaining a common context between the interacting agents is a notoriously hard problem in embodied systems. Such a strict assumptions requires that interacting agents are able to share their perceptions of the environment and simultaneously focus on a particular part of the external world. In essence, the set of registered objects must be identical for both interacting agents. Most recent experiments involving the embodied agents assume the common manual enforcement of shared context settings. In particular, the experimenters align the context of interacting agents by hand [12]. Even though, in practice the strict assumption of LGM can be loosened to a simpler form of topic share restriction, where the agents share only the topic object. Nevertheless, even the simplified restriction is hard to maintain in embodied settings, as it must be guaranteed that the topic is shared between the agents. Usually such a restriction requires sharing of additional (besides the communication itself) information between the interacting agents. As such it is crucial to study the behaviour of the LGM in settings that implicitly allow validation of this basic restriction.

Any deviation from the shared context settings can have twofold interpretation. First, an interaction with non-overlapping contexts represents a case where

the interaction involves an interplay between agents without joint attention, i.e. both focus on different parts of the environment. Second, an interaction may result from an erroneous behaviour of the speaker. For instance, both agents share the current context, however, due to an error the speaker intends to communicate about part of the environment being outside of the current scope. In particular, it is significant to analyse how different learning models, different settings, i.e. population size, context size, and different intensities p_e influence the process of individual semiosis.

3 Model

We extend the original LGM model allowing deviations from the shared context settings. We begin by introducing the notion of system state, i.e. the state of the multi-agent system, as a 7-tuple $S(t)$, as follows:

Definition 1. For each $t \in T = (t_1, \dots, t_{K_T})$ a system state in time point t is a tuple:

$$S(t) = \langle O, W, P_L, P_S, X_O(t), X_P(t), p_e \rangle$$

The system state resembles a general state of the entire multi-agent system in a given point of time. It depicts all of the identifiable objects $O = (o_1, \dots, o_{K_O})$, all of the identifiable words $W = (w_1, \dots, w_{K_W})$, all of the learning agents $P_L = (a_1, \dots, a_{K_{P_L}})$, all of the teaching agents $P_S = (a_1, \dots, a_{K_{P_S}})$.² Additionally it defines the internal processes that govern the dynamics of interaction, i.e. the model of dynamic environment $X_O \subset O$ (available context) and model of agent interaction $X_P \in P_L \times P_S$ (interacting pair). As such, at each discrete time point t the random process X_O models the current state of the environment available to the system. Each interacting agent $(a_L, a_S) \in X_P$ perceives a certain part of its local environment - context $X_O^a(t) \subset O$. Moreover the ability of the system to establish shared context setting between the interacting agents is managed by the shared context intensity $p_e \in [0\%, 100\%]$. It resembles the probability that the teacher managed to successfully establish shared context with the learner, in particular that the selected topic of the utterance is available to the learner in the current context.

3.1 Learner Agent and Teacher Agent

An agent is the most fine-grained autonomous entity present in the system – embodied in the environment and being part of the interacting population. In this research the sole purpose of the agent is to either to learn (Learner Agent) or to teach (Teacher Agent) a language. As such, each agent needs to be equipped with an appropriate semantic infrastructure, allowing it to use the language in order to communicate. As the learner agent is not producing any utterances and

² In the most general case, the set of identifiable objects and the set of all agents in the population can change during the system lifetime, however we assume a simpler case where both the set O, P_L and P_S are finite and static.

is not utilising the selection function it needs to be capable of only registering objects in the environment (autonomous perception) and recording the available linguistic utterance. On the other hand, as the teacher agent is an agent that at the design time is equipped with a static, build-in lexicon there is no need for a learning (lexicon updates) and interpretation. Additionally, the teacher always acts as a speaker and depending on the shared context intensity p_e the speaker either selects an object from the current context η_S (shared context settings) or selects an object outside of the current context η_O (refined context settings).

Each agent is able to perceive objects from the external world $o_i^a \in O$. An object represents a self contained invariant available to agent's perception that is explicitly identified by a unique and strictly internal identifier ($i \sim o_i^a$), i.e., an internal representation of a particular sensorimotor pattern. Further, each agent is able to register and utter words, that are a form of external representations identified by the population as dedicated communication signs. The interconnection of words and objects forms the lexicon \mathcal{L}^a that encapsulates the current state of agent's language. In particular, it is a correlation $\sigma^a(o, w) \in [0, 1]$ between objects $o \in O$ and a words $w \in W$. The higher the association the more definite the agent is that a certain word is an adequate name for an object. Based on the state of the lexicon, simply by traversing the correlations available in the lexicon, the agent is able to produce a name of a particular object (production) and/or interpret a name (interpretation). We incorporate the well established mechanism of interpretation and production [4]. For a given object o the production function ϕ_P^a selects the word w with the maximum correlation ($\phi_P^a(w, L^a(t)) = \text{argmax}_{o_i} \sigma^a(o_i, w)$). Further it is assumed that the predefined lexicon \mathcal{L} is a bijection and that the speaker uniformly samples the current context (shared context settings) or objects being outside of the current context (refined context settings) for topic.

Interaction between agents is the only opportunity for a learner to gain information about the used language. As the interaction is governed by the means of no feedback game routine adopted to the language acquisition task, at each time point $t \in T$ a random pair of agents $(a_L, a_S) \in X_P(t)$ advances in a simple communication. The teacher (a_S) selects a single object $o_T(t) \in X_O^{a_S}(t)$ as the topic [7] and names it $w_T(t) = \phi_P^{a_S}(o_T(t), L^{a_S}(t))$. Next, the uttered word is transmitted to the hearer, that receives it along with the current context $X_O^{a_L}(t)$. Following the idealised settings of LGM the topic is shared among both context ($o_T(t) \in X_O^{a_S}(t) \cap X_O^{a_L}(t)$). However, depending on the intensity p_e , some of the interactions follow the refined routine. In the refined model the teacher selects a topic that is outside of the hearer's current context ($o_T(t) \notin X_O^{a_L}(t)$). Despite the idealised or refined model, the information about current context and uttered name force the hearer to update it's lexicon $L^{a_L}(t) = \psi^{a_L}(w_T(t), X_O^{a_L}(t), L^{a_L}(t-1))$. Further the hearer (a_L) is able to interpret the registered utterance $o_I(t) = \phi^{a_L}(w_T(t), L^{a_L}(t))$ [3]. We should note that without any additional exchange of information the hearer has no means

³ If the intended and interpretation meaning is the same then the game is considered successful.

to predicting whether the current interaction follows the correct (idealised) or incorrect (refined) behaviour of the teacher. Consequently, in the latter case the hearer enforces the incorrect associations.

The agents do not receive any direct feedback concerning the outcomes of the game and the interpreted meaning, i.e., word pair $(w_T(t), o_I(t))$, is regarded as the most probable one. Learning from systematic co-occurrences between words and objects (cross-situational learning) implies that after each interaction the hearer updates its lexicon $L^a(t)$ and modifies the correlations $\sigma^{aL}(o, w_i)$. In general, the update function ψ^{aL} should dampen the correlation $\sigma^{aL}(o, w_i)$ between the received word w_i and currently not perceived objects $o \notin X_O^{aL}(t)$, and enforce the correlations between the received word w_i and currently perceived objects $o \in X_O^{aL}(t)$, while the correlations with other words remain unchanged. In settings involving context with multiple objects, a single interaction is typically insufficient to determine the utilised naming convention, as presumably all objects from the context are equally probable intended meanings. An object can dominate the correlation with a certain word only if it co-occurred with this word more times than with any other object.

We assume that each individual learner is using one of the following alignment strategies [2] – Elimination, Frequency or Damping. *Elimination* strategy assumes the idealised situation, that each exposure of a certain word is accompanied by the presence of the object it denotes. In particular, the true meaning of a particular word lies in the intersection of the different contexts, which it was uttered with. We model the elimination strategy by keeping track of the co-occurrences of word-object (w-o) pairs while eliminating the frequencies of unavailable w-o's. *Frequency* strategy is based on counting the co-occurrences of particular word-object pairs. Each interaction requires the agent to increase the strength of all correlations $\sigma(o, w)$ between the currently available objects and currently uttered word. *Damping* strategy is based on the Frequency strategy, where an additional decrease of co-occurrences between unavailable pairs is performed. Each iteration requires that the agent updates its frequencies, while still damping the co-occurrences of currently uttered word and currently non-available objects.

In order to formulate the dynamics of the alignment process, we identify a major axis of comparison as a numerical summary of the current state of the learning population. We formulate a measure resembling the spread of a naming convention utilised by the teacher among the learning population - reflecting the coherence of names among all existing objects. In particular, it should be stressed that such a measure resembles the observed ability of the system to transfer information from the teacher to the learner. We introduce language coherence μ_{LC} , as the probability that a randomly selected learning agent assign the same name for a randomly selected object from the environment as the teacher agent, as follows:

$$\mu_{LC} = E_{a \in P, o \in O}[\phi_I^a(\phi_P^{aS}(o, L^{aS}), L^a) = o] \quad (1)$$

The lowest possible coherence, i.e. $\mu_{LC} = 0$, reflects a state of no language coherence in the system, as none of the learners uses the same name for any of the objects as the teacher. The highest possible coherence, i.e. $\mu_{LC} = 1$,

represents the state of full coherence, where all learners share teacher's naming convention.

4 Simulation

Using a developed simulation framework⁴ we study the behaviour of the individual semiosis against the refined settings and analyse the behaviour of different learning strategies. Intuitively, a lack of the shared context should result in general inability of the system to reach and maintain an agreement. On the other hand, small disturbances, i.e., only sporadic refinements, should still be marginal and allow the system to reach and maintain a coherent naming convention. Therefore it seems crucial to analyse to what extent the presence of 'invalid' samples affects the behaviour of the system.

First of all we can notice that for high intensity (p_e more than 60%) the language coherence is still able to reach its maximum (after only 3000 iterations Fig. 1). Continuing the decrease of intensity results in continuing decrease of the performance of the alignment process, reaching its final limit for low values of intensity (less than 10%) where the system is unable to align itself. Increasing the number of iterations (15000 Fig. 2) results in the significant increase of both, the agreement and the disagreement area (more than 40% and less than 15%, respectively). As such, it seems obvious that in the limit of number of iterations the system undergoes a rapid phase transition from the agreement state (maximal language coherence) to disagreement state (minimal language coherence).

Focusing on the behaviour of the frequency based learning strategy we test it against different context sizes (Fig. 4). At first, we study the effect of phase transition. For this purpose we test the distributions of final language coherence against different intensities of p_e after 1000, 5000 and 30000 iterations (Fig. 3). Not surprisingly the distinction between the two phases becomes more strict with the increased number of iterations and after 30000 iterations the transition point seems to settle at a value of miscommunication intensity about 21 – 24% (Sec. 5). Secondly, we analyse the behaviour of the system depending on different context settings. As the size of the context determines the indeterminacy of the learner it drastically influences the point of transitions between the phases. In particular, the intuitive behaviour is noticeable as the increase of context size results in drastic decrease of the agreement area, i.e. lowering the transition point (Fig. 3). After 15000 iterations the increase of context size from 2 objects to 3 objects results in nearly a 1/8 decrease in the agreement area, whilst the increase to 4 objects results in nearly a 1/3 decrease. Additionally, we should underline the fact that even with the higher indeterminacy the system managed to cope with miscommunication intensities larger than 50%.

⁴ In the assumed simulations the context sizes are fixed $\forall_{t \in T, a \in P(t)} \|X_O^a(t)\| = c$. All of the presented results are an average of more than 50 consecutive runs demonstrated using boxplot represent the general behaviour of the system. Baseline settings for the experiments include population of 10 Frequency type learners, a single teacher, context size of 2 objects and intensity $p_e = 0$.

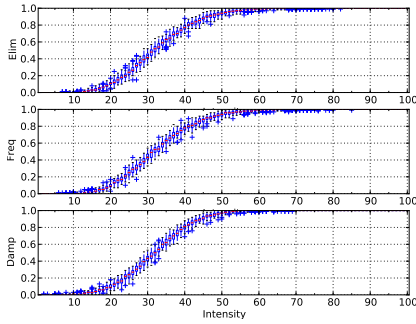


Fig. 1. Language coherence after 3000 iterations against different p_e intensities in case of Elimination(top), Frequency(middle) and Damping(bottom)

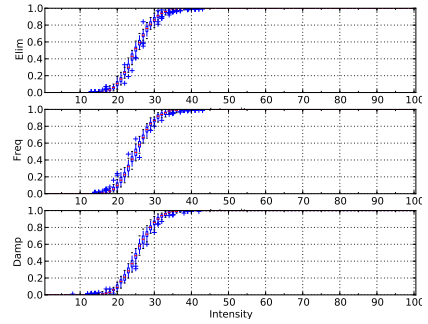


Fig. 2. Language coherence after 15000 iterations against different p_e intensities in case of Elimination (top), Frequency (middle) and Damping (bottom)

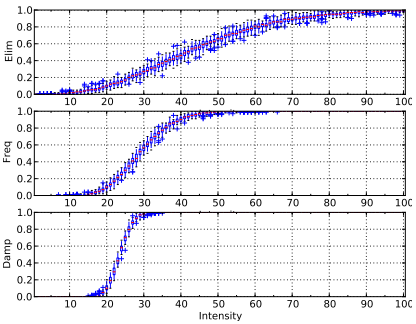


Fig. 3. Language coherence against different intensities after 1000(top), 5000(middle) and 30000(bottom) iterations

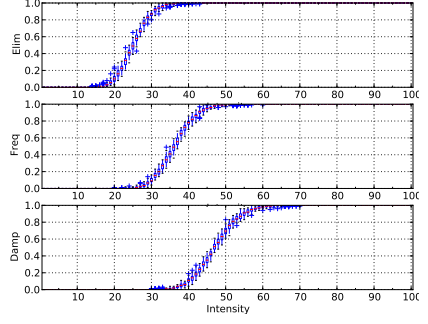


Fig. 4. Language coherence against different intensities for context size of 2(top), 3(middle) and 4(bottom)

5 Analysis and Discussion

The obtained results suggest that despite the distinct learning procedures the influence of changing intensity p_e on the behaviour of the alignment process is similar (despite some minor differences). This fact is due to the similarity on the very basic level of the underlying mechanisms. In particular, all of the proposed learning strategies follow the cross-situational learning stance and all estimate the correct w-o mappings solely based on the registered co-occurrences of particular pairs. Consequently, the influence of the sporadic introduction of 'invalid' samples has a similar effect on all of the tested strategies.

It seems rational that a sporadic introduction of misleading samples should not disrupt the ability of the learner to grasp the language utilised by the teacher. Certainly, as the the invalid samples are not common they can still be treated as

white noise signal that should be filtered out easily. On the other hand, frequent introductions of misleading samples should disrupt the ability of the learner to grasp the teachers convention. Reversing the previous argument, as the invalid samples are frequent then the correct ones tend to be treated as noise and are not properly enforced by the learners. Nevertheless, it seems counter intuitive that the learners are able to grasp the naming convention utilised by the teacher even in a case where only less than one third of interactions follow the shared context settings.

This interesting fact can be easily tracked for the basic type of agent - Frequent Learner (Sec.3.1). Such a learner implicitly assumes that the o-w pair (o, w) with the highest frequency of co-occurrence is the adequate one, whilst all of the other $\forall_{o_i \in O, o_i \neq o}(o_i, w)$ o-w pairs (concerning the word w) are incorrect. This implies that the learner, in order to learn the proper naming convention (the one utilised by the teacher) needs to develop higher frequency values for the correct o-w pairs. The probability that the learner will increase the proper association (o, w) is $P(o, w) = (1 - p_e) \times (p_o^c \times s_o^c)$, where $p_o^c = Pr(o \in X_O(t) \mid \|X_O\| = c)$ is the probability that an object o appears in the current context and $s_o^c = Pr(o = \eta_S(t) \mid o \in X_O(t) \wedge \|X_O(t)\| = c)$ is the probability that object o is selected as topic from the current context. On the other hand, the probability that the learner will increase the association between any other object than o and the word w is $P(o_i, w) = (1 - p_e) \times (p_{o_i, o}^c \times s_o^c) + p_e(p_{o_i, -o}^c \times s_o^{O-c})$, where $p_{o_i, o}^c = Pr(o_i \in X_O(t) \wedge o \in X_O(t) \mid \|X_O(t)\| = c)$ is the probability that both o and o_i appear in the context, $p_{o_i, -o}^c = Pr(o_i \in X_O(t) \wedge o \notin X_O(t) \mid \|X_O\| = c)$ is the probability that object o_i appears in the context where there is no object o and $s_o^{O-c} = Pr(o = \eta_O(t) \mid \|X_O\| = c)$ is the probability of selecting object o as topic from objects outside of the current context. Further, if $P(o, w) > P(o_i, w)$ then on average the correct o-w pairs are more frequent and if $P(o, w) < P(o_i, w)$ then on average the incorrect o-w pairs are less frequent. Consequently, in the former case the learner agent is able to learn the correct convention, whilst in the latter it is unable. These formulas can serve as a criterion for the system of learning agents to successfully align with a teacher agent, i.e. for a population of agents to incorporate a naming convention utilised by the teacher.

The point where the probability of increasing a correct association is equal to the probability of increasing an incorrect association ($P(o, w) = P(o_i, w)$) represents the phase shift. Assuming a system with 10 objects, 10 agents, context size 2, we can estimate the transition to happen roughly around 78% of the miscommunication intensity (due to limited space we omit the proper calculations here). We can notice that the analytically obtained value is perfectly in line with the obtained simulation results (Fig.3).

6 Conclusions

This paper focuses on the studies of the process of individual semiosis in artificial systems. Utilising a simple multi-agent perspective and incorporating the strict language game model we show that a population can agree on a shared

language. We show that in a more realistic settings, where the strict rule of common context is lessened to a case that allows sporadic miscommunication, the interacting population can still agree on a shared naming convention. Most importantly, we have managed to pinpoint the phase transition in the alignment process depending on the intensity of miscommunication in the system. As such, it allows for reasoning about the flexibility of learning strategies. Moreover we have managed to show what are the requirements for the system to maintain the ability to learn the naming convention used by the hearer. Future research will focus on the studies of more complex models of semiosis against the erroneous communication pattern.

Acknowledgments. This paper was partially supported by Grant no. NN519 444939 funded by Polish Ministry of Science and Higher Education (2010-2013).

References

1. Bloom, P.: How children learn the meanings of words. The MIT Press (2002)
2. Blythe, R.A., Smith, K., Smith, A.D.M.: Learning Times for Large Lexicons Through Cross-Situational Learning. *Cognitive Science* 34(4), 620–642 (2010)
3. Smith, K., Smith, A.D.M., Blythe, R.A.: Cross-Situational Learning: An Experimental Study of Word-Learning Mechanisms. *Cognitive Science* 35, 480–498 (2011)
4. DeBeule, J., DeVylder, B., Belpaeme, T.: A cross-situational learning algorithm for damping homonymy in the guessing game. In: *Proceedings of ALIFE X*. MIT Press (2006)
5. Harnad, S.: The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42(1-3), 335–346 (1990)
6. Lorkiewicz, W., Katarzyniak, R.P.: Representing the Meaning of Symbols in Autonomous Agents. In: *Intelligent Information and Database Systems*, pp. 183–189. IEEE (2009)
7. Lorkiewicz, W., Kowalczyk, R., Katarzyniak, R., Vo, B.: On Topic Selection Strategies in Multi-Agent Naming Game. In: *Proceedings of AAMAS 2011*, vol. 2, pp. 499–506 (2011)
8. Saussure, F.: *Course in General Linguistics*, Open Court Pub. Co. (1986)
9. Steels, L.: Language as a complex adaptive system. In: Deb, K., Rudolph, G., Lutton, E., Merelo, J.J., Schoenauer, M., Schwefel, H.-P., Yao, X. (eds.) *PPSN 2000*. LNCS, vol. 1917, pp. 17–26. Springer, Heidelberg (2000)
10. Steels, L.: Language Games for Autonomous Robots. *IEEE Intelligent Systems* (2001)
11. Steels, L.: The Naming Game. *The Electricity Journal* 11(9), 30–33 (2004)
12. Steels, L.: Modeling The Formation of Language in Embodied Agents: Methods and Open Challenges. In: *Evolution of Communication and Language in Embodied Agents*, pp. 223–233. Springer, Heidelberg (2010)
13. Vogt, P.: The emergence of compositional structures in perceptually grounded language games. *Artificial Intelligence* 167(1-2), 206–242 (2005)
14. Vogt, P.: The physical symbol grounding problem. *Cognitive Systems Research* 3(3), 429–457 (2002)

Active Discovery Based Query Federation over the Web of Linked Data

Xuejin Li¹, Zhendong Niu¹, and Chunxia Zhang²

¹ School of Computer Science, Beijing Institute of Technology,
Beijing 100081, China

² School of Software, Beijing Institute of Technology, Beijing, 100081, China
{lixj, zniu, cxzhang}@bit.edu.cn

Abstract. For many applications over the Web of Linked Data, it is a key challenge to access multiple distributed data sources in an integrated and transparent way. In traditional federated query, data from unknown data sources are ignored, which in turn leads to a poor recall of the query result. Furthermore, due to the openness of the Web of Linked Data, it is very difficult to know in advance all data sources in this Web. To overcome these problems we present a novel approach to query the Web of Linked Data. The main idea of our approach is to discover query services that might be used to answer a query during the query execution itself. Our approach can be independently used or the complement of traditional federated query for querying this Web of Data. We provide a prototype implementation and performance evaluation of our work. The experiment shows the feasibility of our approach.

Keywords: Linked Data, Semantic Web, Federated Query.

1 Introduction

In 2006, Tim Berners-Lee outlined four rules [1] for publishing data on the Web that are well known as Linked Data principles. A wide-spread application of these principles enables an evolution of the World Wide Web into a single, globally distributed dataspace -- the Web of Data.

Besides publishing data on the Web according to Linked Data principles, anyone can freely build his Linked Data applications. Data from different data sources can be aggregated and fragmentary information from multiple sources can be integrated to achieve a more complete view. This is not conceivable before. However, for the nature of web, data sources are autonomous and located in different hosts. This makes it difficult to access multiple distributed data sources in an integrated and transparent way. Traditional federated query approaches require knowing in advance potentially relevant data sources. Data from unknown data sources are ignored, which in turn leads to a poor recall of the query result. To provide a general query interface for the Web of Data, it is necessary to obtain and maintain statistics about all data sources in this Web. This is a time consuming work and hardly to be done due to the openness of this Web.

In this paper, we present an approach to query the Web of Linked Data without knowing in advance the data sources that contribute to the query result. The main idea of our approach is to actively discover SPARQL endpoints that can be used to answer a query during the query execution itself. These URIs include URIs in the original query committed by users and those in the intermediate results of the query. Hence, the query is incrementally answered by a growing set of potentially available SPARQL endpoints.

The remainder of this paper is structured as follows. In Section 2 we review related work. Section 3 describes our approach in detail. An evaluation of our approach is given in Section 4. Finally, conclude and discuss future directions in Section 5.

2 Related Work

Olaf Hartig et al. [11] have reviewed different approaches to query the Web of Linked Data. They give an overview of traditional query approaches and novel query approaches and discuss their properties. These approaches can be classified into three categories: data centralization, query federation and link traversal based query. To overcome the problem of ignoring unknown data sources in query federation, they propose an idea combining federated query processing with an active discovery of unknown but potentially relevant linked datasets that are exposed via SPARQL endpoints. However, they do not give an implementation of this idea.

Research on data centralization for querying has a long history in distributed database and search engine research. Findings in distributed database can easily be adjusted for query over the Web of Data. The idea of this kind of approaches is to set up a query service over a collection of Linked Data copied from multiple sources on the Web. OpenLink Software copy data from the majority of the datasets from the Linking Open Data cloud into a central database, referred to as a data warehouse [12]. Data in this central database are exposed via a SPARQL endpoint, and then users can issue SPARQL queries over the whole set of mirrored datasets. While it has fast response times due to no network communication required, implementing this approach might have to overcome some problems discussed by Widom [13]. The similar way in Internet Search has been used to retrieve the Web of Data. Search engines for the Web of Linked Data, such as Sindice [14], Falcons [15] and Swoogle [16], crawl the Web by following RDF links, index discovered data, and provide query interfaces to their indexes. Similar to traditional search engines such as Google and Baidu, they are keyword-based and cannot answer complete SPARQL queries. All of the data centralization approaches suffer from the requirement of much more resources (e.g. storage, compute power, administrative personnel) than is needed to other approaches. Furthermore, huge efforts are needed to maintain the data up-to-date.

Query federation has been a research topic in the area of federated database systems [2] and distributed query processing [17] for a long time. In contrast to data centralization based systems, it distributes the processing of queries to multiple, autonomous sources [2], and, finally, integrates the results of the sub-queries. Benefiting from the ever-increasing amount of data on the Web available at SPARQL endpoints, the idea of federated query is applied to querying the Web of Data. For in stance, DARQ [18] extends the popular query processor Jena ARQ to an engine for federated

SPARQL queries. It provides a service description language which enables the query engine to decompose a query into sub-queries and optimize joins based on predicate selectivities. Another typical federated SPARQL query system is SemWIQ [19] which uses extensive RDF statistics to decompose the original user query. These statistics are generated from a monitoring service. These traditional federated query systems require to know and to select potentially relevant data sources in advance. This requirement prevents a serendipitous discovery and utilization of relevant data from unknown sources [11]. Our approach, due to actively discover relevant data sources during query execution, can overcome this problem.

3 Active Discovery Based Query Federation

This section describes in detail the proposed approach to query the Web of Linked Data by SPARQL. SPARQL [3] is a query language for RDF [4], based on graph patterns and sub-graph matching. The simplest graph pattern defined for SPARQL is the triple pattern which is like the RDF triple except that each of the subject, predicate and object may be a variable. During query evaluation solutions that bind values to the variables are determined. Basic graph pattern BGP is the basic building block from which more complex SPARQL query patterns are constructed. It is a set of triple patterns and can be mixed with value constraints (FILTER) and other graph patterns (OPTIONAL, UNION and GRAPH). In this paper, we discuss the evaluation of graph pattern containing BGP and value constraints and omit other kind of graph patterns.

Our problem is how to answer a query when there is not any priori knowledge about data sources that might be relevant for answering a query. All of the information can be used for answering the query is the query string submitted by users. In this situation, we intend to answer the query by SPARQL endpoints discovered by analyzing the URIs contained in the original query string and the immediate results.

3.1 Service Discovery

Most of data providers publish their data following a set of best practices [5]. Entities are identified by URIs which use `http://` scheme. The `http:` URI scheme implements a two-part approach to identifying resources [6]. It combines a universal distributed naming scheme for owners of resources with a hierarchical syntax for distinguishing resources which share the same owner. Hence, HTTP URIs identifying entities defined by same owners have the same host name and the SPARQL endpoint, if any, is in the same host. Based on this observation, we argue that the SPARQL endpoint of a dataset may be discovered by a URI defined in it.

In this paper, we provide two ways to discover SPARQL endpoints. Possible URLs of SPARQL endpoints are constructed based on URIs in the original query string and in partial results and are verified by sending a simple SPARQL query to it; the responding code being `HTTP_OK` denotes that SPARQL endpoint is available. In addition, we can query whether a data source provides query service by invoking CKAN¹ API. CKAN collected metadata of datasets that have been published in Linked Data format.

¹ <http://ckan.net/>

3.2 Query Planning

The task of query planning is to find relevant data sources and feasible sub-queries. A SPARQL query contains one or more BGPs each containing a set of triple patterns. It may be a URI at the subject, predicate, and object position of a triple pattern. As described in section 3.1, a URI may imply the existence of a SPARQL endpoint. Hence a triple pattern may have zero, one or two relevant data sources according to the third assumption. We wrap a triple pattern as a tuple $TUP=(tp, S, F)$, where tp is a triple pattern, S is the set of SPARQL endpoints that can be used to evaluate tp and F is the set of value constraints that are relevant to tp . The processing of value constraints will be discussed in section 3.3.

Algorithm 1. GenerateSubQueries

```

tups (Input)// a set of TUPs
subqueries (Output)// a set of sub-queries
remains (Output)// the remaining triple patterns
begin
    subqueries :=  $\emptyset$ ;
    remains :=  $\emptyset$ ;
    for each tup  $\in$  tups do
        if tup.S ==  $\emptyset$  then
            remains := remains + {tup}
        else
            if tup.S = {se} then
                sq := subqueries.getQuery(se)
                if sq  $\neq$   $\emptyset$  then
                    sq.BGP := sq.BGP + {tup.t p}
                else
                    subqueries := subqueries +
                        {{tup.tp}, {}, {se}, {}}
                end if
            else
                subqueries := subqueries +
                    {{tp}, {}, tup.S, {}}
            end if
        end if
    end for
    return remains
end.

```

After the step of service discovering, a set of TUPs is created. Triple patterns that have relevant data sources are used to build sub-queries. Each sub-query consists of a set of triple patterns, a set of value constraints, one or two data sources that can answer this sub-query and the intermediate result set. We represent a sub-query as tuple(BGP, F, SES, SLUS), where BGP is a set of triple patterns, F is a set of value

constraints, SES is a set of data sources that can answer the sub-query, and SLUS is the set of solutions of the sub-query. Algorithm 1 shows how the sub-queries are generated. If a triple pattern has exactly one relevant data source the triple pattern will be added to the set of a sub-query for this data source. All triple patterns in this set compose the BGP of the sub-query. If a triple pattern has multiple relevant data sources the triple pattern will individually form a sub-query. The results of this sub-query is the union of results returned by sending a sub-query contains only this triple pattern to all relevant data sources. The remaining part of the query containing triple patterns that no relevant data sources are found will be answered by additional data sources found by URIs in the intermediate results.

3.3 Query Optimization

In distributed query, network latency and bandwidth have the highest influence on query execution time. Thus, the main goal of optimization is to reduce the amount of transferred data and to reduce the number of transmissions, which will lead to less transfer costs and faster query execution.

The evaluation of basic graph patterns and value constraints is order independent. Hence, we can separately consider the set of triple pattern and the set of value constraints in a query. By parsing the original query string, two sets are generated: a set of triple patterns and a set of FILTERs. FILTER expressions that are connected by AND logical operators can be decomposed into sub-FILTERs[7]. Each sub-FILTER will be added into relevant sub-queries. The relevance means that all variables in a sub-FILTER are also used in the triple patterns in a sub-query. If a sub-FILTER is relevant to multiple sub-queries, it will be added into all these sub-queries.

In addition, some variables included in expressions can be eliminated by substituting the value directly in some triple pattern. We use rules based on the results in [9] to rewrite triple patterns in each sub-query.

3.4 Query Execution

For the evaluation of our approach, we provide two ways to execute the query plan: Materialization Query Execution and Semi-materialization Query Execution.

Materialization query execution means that each sub-query is individually evaluated and the results are completely materialized in mediator. The operations that use these intermediate results are executed locally. When no operations need to execute, the final intermediate results will be used to find additional query service that can answer remaining part of the original query. This process will continue to be done until the whole query is answered or no relevant data sources for the remaining triple patterns are found. Hence, to determine results for the whole query we alternately evaluate query parts and find additional query services. During query evaluation, the base storing data sources that provide query service and the base storing data sources that do not provide query service are continuously augmented. This information can be reused during the query execution and in the following query.

Stream processing is another approach to realize query plans. Each solution produced by one operation is passed directly to the operation that uses it [10] and rather small amount of memory is needed compared to approaches that completely materialize intermediate results. In join operation, a sub-query is rewritten with bindings produced by earlier sub-queries. However, this requires the query to be issued repeatedly, once for each variable binding. Due to network latency, it means low query performance. Furthermore, some data servers limit the number of accessing for one client per minute or hour. We combine the stream processing method with the materialization method using semi-joins proposed in [9]. For the smallest network communications, solutions obtained from an earlier sub-query will be grouped and only sent in a batch as conditions in value constraints to their relevant data sources in a join operation.

<pre>SELECT ?drug ?disease ?name WHERE { ?drug drugbank:molecularWeightAverage ?weight. FILTER (?weight <495 && ?weight >494). ?drug drugbank:possibleDiseaseTarget ?disease. ?disease diseasome:name ?name. }</pre> <p style="text-align: center;">Query 1</p>	<pre>SELECT ?disease ?drug ?name WHERE { ?disease diseasome:name ?diseasename. FILTER ?diseasename='Dysprothrombinemia' ?disease diseasome:possibleDrug ?drug. ?drug rdfs:label ?name. }</pre> <p style="text-align: center;">Query 3</p>
<pre>SELECT ?drug ?sideeffect ?label WHERE { ?drug dailymed:name 'Sodium Phosphates'. ?drug owl:sameAs ?drugAlt. ?drugAlt sider:sideEffect ?sideeffect. ?sideeffect rdfs:label ?label. }</pre> <p style="text-align: center;">Query 2</p>	<pre>SELECT ?disease ?drug ?abstract WHERE { ?disease diseasome:name ?diseasename. ?disease diseasome:possibleDrug ?drug. ?drug owl:sameAs ?drugAlt. ?drugAlt dbpedia:abstract ?abstract. FILTER ?diseasename='Dysprothrombinemia' && regex(str(?drugAlt),'http://dbpedia.org/')} }</pre> <p style="text-align: center;">Query 4</p>

Fig. 1. Sample queries

Compared to materialization method, we do not evaluate all sub-queries at a time but compute the count of solutions for each sub-query. In the context of this paper, no statistics are available to do this work. So, we simply send a SPARQL query to relevant data source. If a sub-query has multiple relevant data source, the count of its solutions is the sum of the number of solutions in all relevant data sources. The sub-query that has minimal count of solutions is firstly evaluated and its results are completely materialized. Next, the join operation is executed as follows: Firstly, finding the sub-query that have minimal count of solutions and satisfy the join condition with first sub-query in remaining sub-queries. Secondly, solutions (or called solution mapping) obtained from the earlier sub-query are used to construct value constraints for the found sub-query.

4 Evaluation

In this section we evaluate the performance of the approach presented in this paper. The prototype² was implemented in java based on Jena API and using an extended SPARQL syntax implemented in ARQ2 which allows for aggregate queries. For the sake of reality, no special data servers are set up for our experiments. Example queries are evaluated over real datasets on the Web of Linked data. The information of data sources used in our experiments is shown in Table 1. Our experiments are conducted on a Dual-Core Intel T2300 processor (1.66 GHz), 1GB memory and 1MBit network connection. Our test system runs a professional version of Windows XP with Sun Java 1.6.0_20. The considered depth of paths in service discovering is set to 2 by default.

Table 1. Data Sources

NO.	Description	SPARQL Endpoint
S1	DrugBank	http://www4.wiwiss.fu-berlin.de/drugbank/sparql
S2	Diseasome	http://www4.wiwiss.fu-berlin.de/diseasome/sparql
S3	DailyMed	http://www4.wiwiss.fu-berlin.de/dailymed/sparql
S4	DBpedia	http://dbpedia.org/sparql
S5	SIDER	http://www4.wiwiss.fu-berlin.de/sider/sparql

Table 2. Results for test queries

Query#	Used Data Sources	Solutions
1	S1,S2	44
2	S3,S5	29
3	S1,S2,S3,S4	32
4	S1,S2,S3,S4	32 1

We run each example query in two ways of query execution and evaluated the running time. The queries are shown in Figure 1. It is obvious that these example queries can not be answered using traditional federated SPARQL query approaches [18,19] without knowing in advance relevant data sources that might be relevant for query answering. For all queries we had a timeout of 10 minutes. Query 1 retrieves the name of possible target disease of all drugs with average molecular weight between 494 and 495. Query 2 retrieves all side effects of the drug named Sodium Phosphates. Query 3 retrieves the name of all possible drugs of the disease 'Dysprothrombinemia'. Query 4 retrieves the abstract of all possible drugs of the disease 'Dysprothrombinemia'. Table 2 shows an overview about these queries.

To compare the performance of two ways of query execution we run each example query 6 times and the average execution time is considered for comparisons. Figure 2 shows the query execution time. The experiments show that semi-materialization significantly improves query evaluation performance. Query 1, query 2 and query 3 can

² The source code of our prototype can be downloaded at <http://code.google.com/p/active-discovery-based-query-federation/>

be answered before timeout. For query 4, the materialization query executions took longer than 10 min to answer and timed out. This is due to the fact that sub-queries in query 1, query 2 and query 3 have smaller sets of intermediate results than those in query 4. The execution time of the semi-materialization query executions of four example queries is quite reasonable.

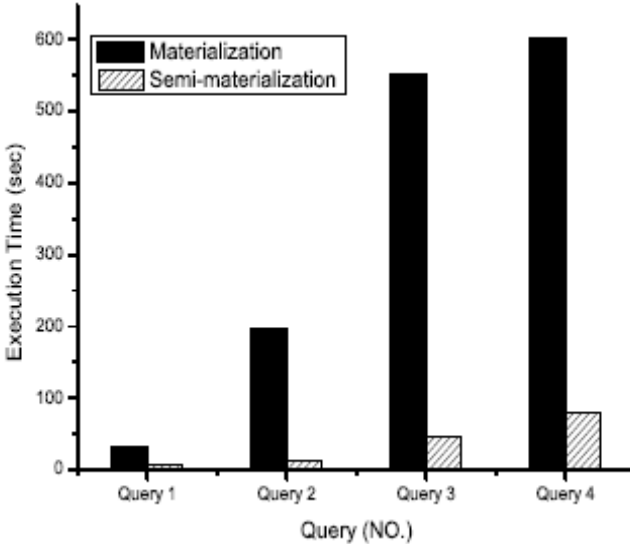


Fig. 2. Query execution times

Our evaluations show that semi-materialization query execution is more performant than materialization query execution. All queries are answered within less than one minute. Of course, materialization query execution may be in high performance if the size of immediate results is small, but it is common that the size of immediate results is very large when no any value constraints exist in real-world queries. The smaller the size of immediate results in a query is, the closer the performance of these two ways of query execution is.

5 Conclusion and Future Work

In this paper we have introduced a way to query the Web of Linked Data. The concept of Active Discovery based Query Federation is firstly proposed in [11]. Our approach is based on analyzing URIs in a triple pattern to discover SPARQL endpoints that can be used to answer a query during the query execution itself. The query plan is executed by two ways: Materialization Query Execution and Semi-Materialization Query Execution. Our experiments on real-world environment show that the later is in faster query performance.

A set of best practices have been absorbed by more and more data providers on the Web of Linked Data. Our approach benefits from the relevance between the URIs

identifying entities defined in one dataset and the URL of SPARQL endpoints it provides. In addition to this relevance, a high number of links in the Web of Linked Data is significant for the more complete results. The more links between data sources the more query services can be discovered. In our approach, only logical optimization can be used before query execution because of the lacking of statistics about relevant data sources. Hence, we are investigating possibilities to combine traditional federated SPARQL query with our approach. Such a combination poses new challenges to query planning and optimization. We are currently working on these issues.

In federation query, query service is necessary for relevant data sources. However, providing a SPARQL endpoint is not required in the Linked Data principles. According to CNKA, there are 168 endpoints out of 202 data sources at the time of writing this paper. Both traditional federation query and our approach just omit those dataset not providing query services. For a more general query interface, additional technologies should be considered. The link traversal based query execution is a possible solution. In the future work, we propose to combine these technologies into a hybrid one.

References

1. Berners-Lee, T.: Design Issues: Linked Data. Online, <http://www.w3.org/DesignIssues/LinkedData.html> (retrieved May 25, 2009)
2. Sheth, A.P., Larson, J.A.: Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Comput. Surv.* 22(3), 183–236 (1990)
3. Prud'hommeaux, E., Seaborne, A.: SPARQL Query Language for RDF. W3C Recommendation (January 2008), <http://www.w3.org/TR/rdf-sparql-query/>
4. Manola, F., Miller, E.: RDF Primer, W3C Recommendation (2004), <http://www.w3.org/TR/rdf-primer/>
5. Heath, T., Hausenblas, M., Bizer, C., Cyganiak, R.: How to publish linked data on the web (tutorial). In: *Proc. 7th Int. Semantic Web Conf.* (2008)
6. Thompson, H.S., Orchard, D.: URNs, Namespaces and Registries. W3C TAG Finding (August 2006), <http://www.w3.org/2001/tag/doc/URNsAndRegistries-50.html>
7. Bernstein, A., Christoph Kiefer, M.S.: OptARQ: A SPARQL Optimization Approach based on Triple Pattern Selectivity Estimation. Technical Report 2007.03, Department of Informatics, University of Zurich (2007)
8. Pérez, J., Arenas, M., Gutierrez, C.: Semantics and complexity of SPARQL. In: Cruz, I., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L.M. (eds.) *ISWC 2006. LNCS*, vol. 4273, pp. 30–43. Springer, Heidelberg (2006)
9. Zemanek, J., Schenk, S., Svatek, V.: Optimizing sparql queries over disparate rdf data sources through distributed semi-joins. In: *ISWC 2008 Poster and Demo Session Proceedings*, CEUR-WS (2008)
10. Garcia-Molina, H., Widom, J., Ullman, J.D.: *Database Systems: The Complete Book*. Prentice-Hall, Inc., Upper Saddle River (2002)
11. Hartig, O., Langegger, A.: A Database Perspective on Consuming Linked Data on the Web. *Datenbank-Spektrum* 10(2) (2010)
12. Chaudhuri, S., Dayal, U.: An overview of data warehousing and olap technology. *ACM Sigmod Record* 26(1), 65–74 (1997)

13. Widom, J.: Research problems in data warehousing. In: Proceedings of the Fourth International Conference on Information and Knowledge Management, pp. 25–30. ACM (1995)
14. Oren, E., Delbru, R., Catasta, M., Cyganiak, R., Stenzhorn, H., Tummarello, G.: Sindice.com: A document-oriented lookup index for open linked data. *International Journal of Metadata, Semantics and Ontologies* 3(1) (2008)
15. Cheng, G., Qu, Y.: Searching linked objects with falcons: Approach, implementation and evaluation. *International Journal on Semantic Web and Information Systems (IJSWIS), Special Issue on Linked Data* 5(3), 49–70 (2009)
16. Ding, L., Finin, T.W., Joshi, A., Pan, R., Cost, R.S., Peng, Y., Reddivari, P., Doshi, V., Sachs, J.: Swoogle: A search and metadata engine for the semantic web. In: Proceedings of the 13th ACM Conference on Information and Knowledge Management (CIKM), pp. 652–659 (November 2004)
17. Kossman, D.: The State of the Art in Distributed Query Processing. *ACM Computing Surveys* 32(4), 422–469 (2000)
18. Quilitz, B., Leser, U.: Querying Distributed RDF Data Sources with SPARQL. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) *ESWC 2008*. LNCS, vol. 5021, pp. 524–538. Springer, Heidelberg (2008)
19. Langegger, A., Wöb, W., Blöchl, M.: A Semantic Web Middleware for Virtual Data Integration on the Web. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) *ESWC 2008*. LNCS, vol. 5021, pp. 493–507. Springer, Heidelberg (2008)

A Data Imputation Method with Support Vector Machines for Activity-Based Transportation Models

Banghua Yang^{1,2}, Davy Janssens¹, Da Ruan^{1,3}, Mario Cools¹,
Tom Bellemans¹, and Geert Wets¹

¹IMOB, Hasselt University, 3590, Hasselt, Belgium

²Department of Automation, College of Mechatronics Engineering and Automation;
Key Laboratory of Power Station Automation Technology;
Shanghai University, Shanghai 200072, China

³Belgium and Belgian Nuclear Research Centre (SCK-CEN), Belgium
{Banghua.yang, Davy.janssens, Mario.cools, tom.bellemans,
geert.wets}@uhasselt.be, yangbanghua@shu.edu.cn

Abstract. In this paper, a data imputation method with a Support Vector Machine (SVM) is proposed to solve the issue of missing data in activity-based diaries. Here two SVM models are established to predict the missing elements of ‘number of cars’ and ‘driver license’. The inputs of the former SVM model include five variables (Household composition, household income, Age oldest household member, Children age class and Number of household members). The inputs of the latter SVM model include three variables (personal age, work status and gender). The SVM models to predict the ‘number of cars’ and ‘driver license’ can achieve accuracies of 69% and 83% respectively. The initial experimental results show that missing elements of observed activity diaries can be accurately inferred by relating different pieces of information. Therefore, the proposed SVM data imputation method serves as an effective data imputation method in the case of missing information.

Keywords: Activity-based transportation models, Support Vector Machine (SVM), Data imputation, Missing data.

1 Introduction

Activity-based approaches in transportation models aim at predicting which activities are conducted where, when, for how long, with whom, the transport mode involved and so on. The activity-based approach is a sound option to model people’s travel behavior, which has set the standard for travel demand modeling during the last decade [1]. The basic premise of this approach is that travel demand is derived from the activities that individuals and households need or wish to perform. A dynamic activity-based travel demand framework, FEATHERS (Forecasting Evolutionary Activity Travel of Households and their Environmental RepercussionS) has been developed for Flanders (the Dutch speaking region of Belgium) based on the above aim [2-4]. The FEATHERS to be applied for the whole Belgium is currently being developed at IMOB. To build the FEATHERS model that can predict all of those above facets, one

requires data on all these facets. Clearly, the data collection is a huge challenge [5]. One of the main data sources is activity-based diaries. The collection of diary data frequently brings a huge burden on respondents. Consequently, activities diaries tend to contain incomplete information due to various reasons, which is a serious problem because activity-based models require complete diary information [6].

Activity diaries used in the existing FEATHERS mainly contain individual surveys and household surveys. The collected surveys are composed of 8551 samples of households and 15888 samples of personals in Belgium. Each sample includes many variables. Among all samples, about 10% samples of households and 5% samples of personals exist missing information. If all samples that contain any missing values are deleted and the analysis is then carried out on the samples that remain, some serious drawbacks will be brought. One of drawbacks is the reduction of samples, which will affect the predicting reliability and quality of the FEATHERS model. The other is that the elimination of useful information in the sample will result in serious biases if the samples are not missing completely. The interest of this paper has centered on performing data imputation, the process by which missing values in a data set are estimated by appropriately computed values, thus constructing a complete data set.

More recently, with the development of computer science and technology, some artificial intelligence and machine learning techniques have arisen in the area of missing data treatment, such as neural networks, fuzzy logic systems, and rough sets, which stimulate the missing data research to a new stage. In this paper, the support vector machine (SVM) is proposed to predict the missing values of two variables. The SVM is a new generation learning system based on recent advances in statistical learning theory [7, 8]. SVMs have been applied in many areas such as text categorization, hand-written character recognition, image classification, and bio-sequences analysis [9, 10]. Here two SVM models are established to predict the missing elements of number of car and driver license respectively. The first SVM model to predict the ‘number of car’ achieves an accuracy of 69%. Meanwhile, the second SVM model to predict ‘driver license’ can obtain an accuracy of 83%. The results are verified by a four-fold cross-validation.

2 Data Description

Since the objective of the activity diary is to give a representative description of the travel behavior of the population in Belgium, the target population in the project was defined as “all the people residing in Belgium, regardless of their place of birth, nationality of any other characteristics”. Activity diaries used in the FEATHERS mainly contain individual surveys and household surveys. The data were collected in 2010, in which the individual surveys were carried out among Belgian citizens aged 6 years and above. The total number of collected sample equals 8551 households comprising 15888 individuals. In the activity diary, a household record has many variables, such as 1) the name, sex, nationality, educational certificate, and professional status of each household member; 2) the type of vehicle, number of the specified vehicle, and purchase year of the specified vehicle that the household possess; 3) the place of residence, net income of the household, etc. The individual survey includes person ID,

the mode of travel, number of trips, the start time, the arrival time, activity type, activity duration, activity location ID, and the driver license.

All collected surveys are summarized to excel tables. Table 1 shows a small part of samples from household surveys and Table 2 shows ones from individual surveys, in which the shading cells represent the missing data. It should be pointed out that the title of each column in tables is only a simple nomination or the question number. All data should be preprocessed to meet requirements of the FEATHERS. The FEATHERS need five files (Households, Persons, Activities, Journeys and Lags). Each file includes many variables respectively. Here the interest centers on the Households and Persons. There are eight variables (HouseholdID, Household locationID, Household composition, Socio-economic class, Age oldest household member, Children age class, Number of cars and Number of household members) in the Households file. Meanwhile, the Persons file includes six variables (PersonID, HouseholdID, Personage, Work status, gender and Driver's license). Among all samples, about 10% samples of households and 5% samples of personals exist missing information. How to estimate the missing information and so improve the number and quality of samples is the main concern of this paper. The following section will describe a SVM method to process the missing information.

Table 1. Households data

H_id	H_nper	H_oldjour	H_bir	H_sex	H_pos	H_status	H_q2	H_q1
100015	2	1/12/2009	1983	M	2	9	1	8
100048	2	4/12/2009	1965	F		7	0	3
100081	1	7/12/2009	1941	M	1	5	1	3
100103	1	9/12/2009	1954	M	1	9		3
100125	1	11/12/2009	1937	F	1	5	0	99
100235	2		1942	M	1	5	2	7
...
100466	1	17/12/2009	1971	F	1	4	0	2

Table 2. Individuals data

Num	I_refday	I_nbtrip	I_q1a	I_q1b	T_hd	T_mind	...I_q4a	I_q4a
10001	1/12/2009	3	2	3	06	30	...	3 1
10001	1/12/2009		2	2	06	55	...	3 1
10004	4/12/2009	2	1	2	09	00	...	2
10009	9/12/2009	9	1	5	07	00	...	3 1
10011	9/12/2009	0	1		16	00	...	2
...	1
10017	17/12/2009	2	2	2	11	15	...	0 1

3 Data Imputation Method with SVM

3.1 The Basic Theory of SVM

A SVM is one of supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. It takes a set of input data and predicts which of two possible classes each given input belongs. The SVM performs classification by constructing a hyperplane that optimally separates the data into two categories. New examples are then mapped into that same space and predicted to belong to a category based on which side of the hyperplane they fall on. A good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class (so-called margin), since in general the larger the margin the lower the generalization error of the classifier. So the key technique to SVMs is how to obtain a hyperplane that has the largest margin.

For a given training data set D ,

$$D = \{(x_i, y_i) \mid x_i \in R^p, y_i \in \{-1, 1\}, i = 1 \dots n\}; \tag{1}$$

where the y_i is either 1 or -1, indicating the class to which the point x_i belongs. Each x_i is a p -dimensional real vector. The classification of SVM is to find the maximum-margin hyperplane that divides the points having $y_i = 1$ from those having $y_i = -1$. The hyperplane with the maximum-margin can be described as follows:

$$w \bullet x - b = 1 \quad \text{and} \quad w \bullet x - b = -1 \tag{2}$$

where \bullet denotes the dot product and w the normal vector, $\frac{b}{\|w\|}$ determines the offset of the hyperplane. If all data points are prevented from falling into the margin, a constraint can be added as follows:

$$y_i(w \bullet x_i - b) \geq 1, \text{ for } i = 1 \dots n; \tag{3}$$

The problem about the maximum-margin hyperplane can be described as:

$$\text{Minimize } \|w\|, \text{ subject to } y_i(w \bullet x_i - b) \geq 1, \text{ for } i = 1 \dots n; \tag{4}$$

At last, the problem can then be expressed by means of non-negative Lagrange multipliers α_i as [8]

$$\min_{w,b} \max_{\alpha} \left\{ \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i(w \bullet x_i - b) - 1] \right\} \tag{5}$$

If there does not exist any hyperplane that can split the "-1" and "1" examples, a soft margin method can be introduced to choose a hyperplane that splits the examples as cleanly as possible. The method introduces a slack variable, ξ_i , which measures the degree of misclassification of the data x_i ,

$$y_i(w \bullet x_i - b) \geq 1 - \xi_i, \text{ for } i = 1 \dots n; \tag{6}$$

At the same time, the optimization problem becomes:

$$\min_{w, \xi, b} \max_{\alpha, \beta} \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i [y_i (w \bullet x_i - b) - 1 + \xi_i] - \sum_{i=1}^n \beta_i \xi_i \right\}, \quad \alpha_i, \beta_i \geq 0 \quad (7)$$

In general, an original problem may be stated in a finite dimensional space. The maximum-margin hyperplane needs to be obtained by transforming the original space to a high-dimensional space, which is proposed to make the separation easier in that space. To keep the computational load reasonable, the SVM scheme is designed to ensure that dot products may be computed easily by the kernel function in the original space. This problem can now be solved by standard quadratic programming techniques and programs. The detailed description about SVM can be found in [9].

3.2 Prediction for ‘Number of Cars’

As described in Section 2, there are eight variables that are needed in the Households file to meet the requirement of the FEATHER. However, about 8% samples in this file miss information of ‘number of cars’. Since different variables are interrelated, such as ‘number of cars’ should be dependent on the income of a household, the missing variable information in one sample is expected to be inferred by other variables in the same sample. Here the missing ‘number of cars’ are predicted by other five variables. The SVM model to predict ‘number of cars’ is shown in Figure 1. The inputs of the SVM model include Household composition, household income, Age oldest household member, Children age class and Number of household members. The output of the SVM is the ‘number of cars’, which has three values (0, 1, 2). These three values represent that a household has no car, one car, two cars or more, respectively.

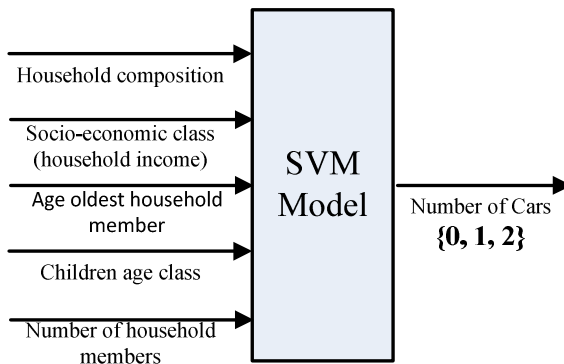


Fig. 1. The SVM model to predict ‘number of cars’

Obviously, the prediction is a three-category classification. However, a SVM is often used to distinguish two categories. The approach used in this paper for doing so is to reduce the single multiclass problem into multiple binary classification problems. First, a strategy called “one against many” is adopted, in which each category is split

out and all of the other categories are merged. According to this strategy, three SVM models (SVM Model 1, SVM Model 2, and SVM Model 3) are established, which is shown in Fig.2. The SVM model 1 is used to distinguish label 0 and the rest, where label 1 and label 2 are merged. The SVM model 2 is used to distinguish label 1 and the rest, where label 0 and label 2 are merged. The SVM model 3 is used to distinguish label 2 and the rest, where label 0 and label 1 are merged. The classification accuracy is obtained by means of a four-fold cross-validation as shown in Fig.3. From Figure 3 the SVM model 1 obtains the highest classification accuracy. And so the SVM model 1 is selected to distinguish label 0 and the rest. If a new sample is assigned as label 0, the classification is over. Otherwise, the sample needs to be distinguished continuously between 1 and 2, which is shown in Fig.4. After the classification of the SVM model 1 and the SVM model 4, the final label will be decided. The SVM model to predict the ‘number of cars’ can achieve an accuracy of 69% by means of a four-fold cross-validation.

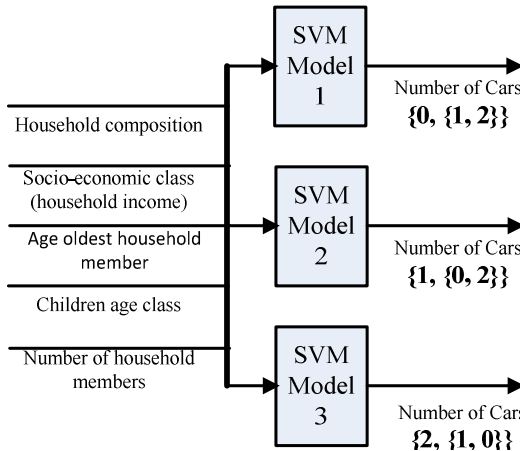


Fig. 2. SVM models based on “one against many” strategy

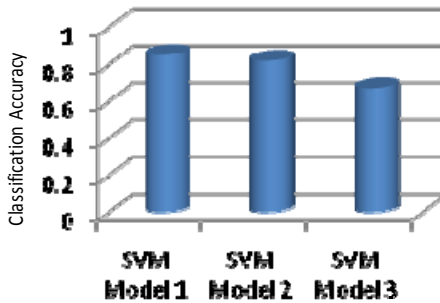


Fig. 3. Classification accuracies of different SVM models

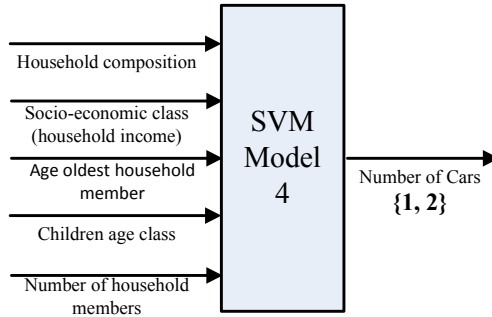


Fig. 4. The SVM model between label 1 and label 2

3.3 Prediction for ‘Driver License’

As described in Section 2, the Persons file used in the FEATHERS include six variables (PersonID, HouseholdID, Personage, Work status, gender and Driver’s license). However, about 2% samples in this file miss information of ‘driver license’. Considering the relativity between Personage, Work status, gender and driver license, here the missing ‘driver license’ data are predicted by the other three variables (Personage, Work status, and gender). The SVM model to predict ‘driver license’ is shown in Figure 5. The inputs of the SVM model include Personage, Work status and gender. The output of the SVM is the ‘driver license’, which has two values (0 and 1). These two values represent that a person has no driving license or has a driving license. The SVM model to predict the presence of the driving license can obtain an accuracy of 83 % by means of a four-fold cross-validation.

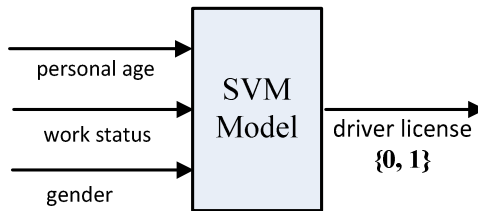


Fig. 5. The SVM model to predict ‘driving license’

4 Results and Discussions

In this paper, two SVM models are established to predict the ‘number of cars’ and ‘driver license’ respectively. During the prediction for ‘number of cars’, two necessary steps are taken. In the first step, a strategy called “one against many” is adopted and the best result is obtained by distinguishing between label 0 and the rest. The label 0 represents a household has no car and the rest (label 1 and label 2) represents a household has one car, two cars or more. The obtained result shows that a household without any car can be best differentiated from the one with cars, which is concordant

with our intuitive impression. In our real life, a household without any car has many differences with the one with cars according to our feeling. We also established models for any two categories, that is between label 0 and label 1, label 0 and label 2, labels 1 and label 2. The classification results of these two categories are 78%, 90% and 79% respectively, which show that a household without any car can be distinguished from the one with two cars better than other two instances. All these results show that the established models are consistent with the real life. If a new sample is assigned 'the rest' in the first step, it will then be assigned to either label 1 or label 2 in the second step.

In the prediction of 'number of cars' and 'driver license', the accuracies of 69% and 83% are obtained respectively by means of a four-fold cross-validation. The missing information in these two variables is imputed by related known information in other variables. The prediction results keep identical with analyzed facts, which show that the established SVM models are feasible to complete information in the FEATHERS model. Of course, how to select some optimized and related variables and how to optimize some parameters of the SVM models will be studied further.

5 Conclusions

The establishment of the FEATHERS model needs large amount of complete data. How to improve the quality and quantity of sample data under existing activity-based diaries will be one of problems to be solved in the FEATHERS. Aiming at this problem, a data imputation method based on SVM is proposed. Two SVM models are established to predict the missing information of variables called 'number of cars' and 'driving license' using related other variables. The prediction accuracies of 69% and 83% are obtained respectively by means of a four-fold cross-validation. The initial results show the feasibility of the proposed method. The whole samples can be increased 1%-8% after an accurate prediction. The established SVM models can be used to new samples, which provide a good approach to perfect the missing data. The further research is about how to improve the prediction accuracy by a better imputation method.

Acknowledgments. The project is a post-doctoral research grant supported by BELSPO, the Belgium Science Policy and with co-financing from the European commission.

References

1. Moons, E., Wets, G.: Tackling Non-response in Household Travel Surveys: A Case Study. In: Proceeding of the 56th Session of the International Statistics Institute, Lisbon, CD-ROM (2007)
2. Janssens, D., Wets, G., Timmemans, H.J.P.: Modeling Short-term Dynamics in Activity-Travel Patterns: the Feathers Model. In: Proceedings of the 11th World Conference on Transportation Research. WCTRS, Berkeley, California, USA, CD-ROM (2007)

3. Bellemans, T., Kochan, B., Janssens, D., Wets, G.: Implementation Framework and Development Trajectory of FEATHERS Activity-based Simulation Platform. Transportation Research Record: Journal of the Transportation Research Board, No. 2175, Transportation Research Board of the National Academies, Washington, D.C., 111–119 (2010)
4. Cools, M., Moons, E., Bellemans, T., Janssens, D., Wets, G.: Surveying Activity-travel Behavior in Flanders: Assessing the Impact of the Survey Design. In: Proceedings of the BIVEC-GIBET Transport Research Day, Part II, Brussels, pp. 727–741 (2009)
5. Nakamya, J., Moons, E., Koelet, S., Wets, G.: Impact of Data Integration on Some Important Travel Behavior Indicators. Transportation Research Record: Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C, 89–94 (2007)
6. Arentze, T., Hofman, F., Nelly Kalfs, N., Immermans, H.: System for Logical Verification and Inference of Activity (SYLVIA) Diaries. Transportation Research Record 1660, 156–163 (1999)
7. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer (1995)
8. Cortes, C., Vapnik, V.: Support-vector Networks. Machine Learning 20, 273–297 (1995)
9. Hearst, M.A., Schölkopf, B., Dumais, S., Osuna, E.: Trends and Controversies - Support Vector Machines. IEEE Intelligent Systems 13, 18–28 (1998)
10. Sanchez, V.D.: Advanced Support Vector Machines and Kernel Methods. Neuro Computing 55, 5–20 (2003)

The Study on Integer Overflow Vulnerability Detection in Binary Executables Based Upon Genetic Algorithm

Baojiang Cui¹, Xiaobing Liang¹, and Jianxin Wang²

¹ School of Computer, Beijing University of Posts & Telecommunications,
Beijing 100876, China

² School of Information Science & Technology, Beijing Forestry University,
Beijing 100083, China

{cuibj, liangxb}@bupt.edu.cn, wangjx@bjfu.edu.cn

Abstract. The automatic identification of security vulnerabilities in the binary code is still a young but important research area for the security researchers. In recent years, the number of identified integer overflow vulnerabilities has been increasing rapidly. In this paper, we present a smart software vulnerability detection technology, which is used for the identification of integer overflow vulnerabilities in the binary executables. The proposed algorithm is combined with debugger module, static analysis module and genetic algorithm module. We use the fitness function to guide the generation of the tested data and use static analysis to provide the information that the genetic module needs. Theory analyses and experiment results indicate that the detection technology based upon genetic algorithm can identify the exceptions in the object program and is more efficient than the common Fuzzing technology.

Keywords: Fuzzing technology, Integer overflows vulnerability, Binary executables, Static analysis, Genetic algorithm.

1 Introduction

Recently, the number of identified integer overflow vulnerabilities has been increasing rapidly. According to the 2006 CVE report [1], integer overflows rose to the second most common vulnerability in the operating system vendor advisories. Integer overflows are an underestimated source of vulnerabilities. The overflow itself is usually not exploitable, but it can lead to other classes of vulnerabilities, frequently buffer overflows. Furthermore, a number of integer overflow attacks have been witnessed in recent years, such as the integer overflow vulnerabilities in Adobe PDF [2] and CamlImages TIFF [3].

Since integers typically have a fixed size (e.g., 32bits) for particular machine architecture, they can store fixed maximum value. Operations that result in a value greater than this maximum can cause the value to wrap-around: this well-known condition is called overflow. Four types of integer overflow vulnerabilities are described in [4], such as overflow underflow, signedness error and truncation.

Fuzzing[5] is a very simple technique of black-box testing. The basic idea is to attach the inputs of a program to a source of random data: any resulting failure is taken

as a symptom of a defect in the program. This approach was successfully used to discover vulnerabilities in the object programs, though released without source code [5,6]. Since its introduction, the Fuzzing approach has become a major tool in security research area. However, while this approach works well when aimed at searching genetic failures, security researchers are usually interested in more specialized corner cases, which can crash the object binary code. By using random data, even a single equality condition on just one byte of the input data could be difficult to satisfy wittingly: in fact, satisfying randomly a 32-bit equality corresponds to guessing by chance the correct value out of four billion possibilities. So, the efficiency of the common Fuzzing technology is very low.

As software becomes larger and more complex, the number of condition branches and loops in the software indicates an inevitable rising. The complex condition branches and deeper nested loops in the software result in the serious problem of path blowup. Furthermore, the input scope of a program is very huge, while the subset of the input space which a malicious user can exploit to drive it into an insecure state is very small. All the above reasons that are described make the efficiency of the Fuzzing technology more slow than before. Genetic Algorithms have been widely used in many areas to search for optimal designs, and they have shown their worth in solving problems where normal algorithms are too inefficient or ineffective [7]. They have the property that under most circumstances GAs will found a good local optimum. Genetic algorithm has been used to generate input data for software test. Given the program source code, [8] presents a technique based on Genetic Algorithm for automatically generating test-data to test exceptions.[9] use Genetic Algorithm to detect buffer overflow via automatic test input data generation. However, as source code is not always available, the state-of-the-art techniques have to rely on dynamically running the software, exploring program paths. [10] presents an extension of traditional “black box” fuzz testing using a genetic algorithm based upon a Dynamic Markov Model fitness heuristic.

In this paper, we present a smart software vulnerability detection technology, which is used for the identification of integer overflow vulnerabilities in the binary executables. We do not directly analyze the assembly code because the semantics of assembly instruction is very complicated. Unlike common fuzz technique, we add quality assessment and metric on the generation test cases that are used for testing the binary programs. Our method uses GA’s fitness function to guide the generation of input data that the detection technology needs, which can improve the efficiency of GA. In order to guide the test data generation more effectively, we develop a multi-objective fitness function to guide the generation of the test data.

The rest of the paper is organized as follows: Section 2 describes the design of the whole system. Section 3 provides an experimental evaluation of our technique on a small program code. Finally, Section 4 concludes.

2 System Design

In this section, the detailed design and function of the three different modules in our system will be described.

2.1 The Debugger Module

The debugger module is the basic platform of our prototype system, which monitors the execution of the binary executables and records the related execution information, such as the execution path resulted by one input and the exceptions caused by one input. The major functions of the debugger include:

- The debugger identifies and intercepts the bottom API functions of Windows system in order to obtain the input spot of the binary executable.
- The debugger can read the internal data of the binary executables, such as the data in the register, heap and stack and so on.
- The debugger tracks the binary executables, computes and updates the variable values in the register and memory position of the binary.

According the information that the control flow extracting module provides and the help of the debugger module, we can record the execution trace of an input, compute the basic block coverage ratio, the least probable execution paths and the number of danger function execution. All the information that is obtained in the debugger module will send to the genetic algorithm module to dynamically compute the fitness function value of an input.

2.2 The IDA Static Analysis Module

In this paper, we make use of IDA pro [11] to disassemble the binary executables. IDA pro can resolve the executables with the multiple formats, including Portable Executable, Executable and Linking Format and other formats. IDA pro provides a lot of useful information: procedure boundaries and calls to library functions.

Control flow structure and basic blocks	Danger function list	Function list
Head:01001348	010010fc lstrcpyW	01001000 RegQueryValueExW
Tail:01001937	01001118 lstrcatW	01001004 RegCloseKey
Exit1:0100193c	01001130 lstrcpyW	01001008 RegCreateKeyW
Exit2:01001939	01001280 wsprintfW	0100100c IsTextUnicode
JxxA:01001937	01001314 _snwprintf	01001010 RegQueryValueExA

Fig. 1. The partial analysis result of notepad.exe

We develop an IDA static analysis tool to extract the control flow structure, the basic blocks and the using information about the danger functions of the binary executable. Through the analysis of jump instructions, we can obtain the basic blocks and control flow structure of the binary executable. Using the library functions provided by the IDA pro, the danger function list can be got. We use the IDA static analysis tool to analyze the notepad.exe, the partial experiment results are shown in the Figure 1.

2.3 The Genetic Algorithm Module

Genetic algorithms were developed initially by Holland et al. in the 1960s and 1970s [12]. They attempt to model the natural genetic evolutionary process. Generally speaking, a GA may be defined as an iterative procedure that searches for the best solution of a given problem among an evolving population, represented by a finite string of symbols, the genome. The search starts from an initial population of individuals, often randomly generated. At each evolutionary step, individuals are evaluated using a fitness function. High-fitness individuals will have the highest probability to reproduce. Successive populations are evolved using the genetic operations of crossover and mutation. The aim is that through the use of genetic operations the population will converge towards a global solution.

2.3.1 Crossover and Mutation

In this paper, we aim to detect the vulnerabilities in the windows application programs whose input style are reading files. The crossover and mutation operations are focus on the bytes in the input file.

The operation of crossover make two chromosomes exchange a part of each one, with two new chromosomes (of the next generation) produced. Two chromosomes of one generation can mutate and produce another two chromosomes of the next generation. The operations of crossover and mutation are illustrated in figure 2.

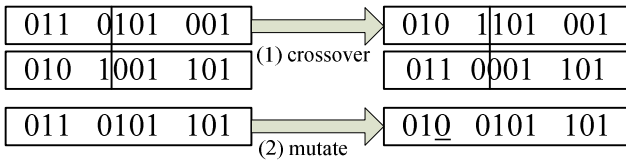


Fig. 2. Crossover and mutate

2.3.2 Fitness Function

A control flow graph for a binary executables is a directed graph with nodes corresponding to blocks of sequential instructions and edges corresponding to non-sequential instructions that join the basic blocks (i.e. condition branch instructions). It can be described as follows: $G = (N, E, s, e)$, where N is a set of nodes, E is a set of edges of the form (n_i, n_j) , s and e are unique entry and exit nodes such that $s, e \in N$. A node, $n \in N$, is a sequential instructions such that if any instruction of the block is executed, then all are executed. This is also known a basic-block. An edge $(n_i, n_j) \in E$ represents a possible transfer of control from the basic block n_i to the basic block n_j . For branch instructions the edges are associated with a branch predicate. This describes the conditions which must hold for the branch to be taken.

In order to better guide the search toward regions of the input parameters where the integer overflows are likely to be exposed. Some possible elements to include are the following:

- basic block coverage: high basic block coverage contributes to increasing the likelihood of an exception or an integer overflow.
- the least probable execution paths coverage: the higher this coverage, the higher path coverage and the higher likelihood of in integer overflow. We will reward those genomes in the population that represent input that takes previously unexplored execution paths as well as rare (i.e. difficult) execution paths. In order to distinguish the paths that have been executed by the test cases, we introduce a hash function which is constructed by the basic block information that the path has exercised. We maintain a hash table during the iteration process of the GA. If the hash value of a path is same as the value stored in the hash table, we think that the path has been exercised and give the test case a lower fitness value.
- number of executions of danger functions or danger operations: the higher the number, the higher the likelihood of generating exceptions and integer overflows;

The basic block coverage gives the percentage of the executed blocks; the least probable execution paths show that we expect to execute the unexplored paths as well as rare execution paths. Finally, the number of danger functions or danger operations is the number of time danger functions or danger operations are executed given test data. If the tainted data is delivered to danger functions or danger operations, we think that the binary executables may result in the integer overflows. The fitness function used in this paper is composed as

$$Fitness(x) = w_1 * bbc_x + w_2 * l cov_x + w_3 * rIndex * \log(D_x) ,$$

Where for each individual x :

- bbc_x is the basic block coverage ratio;
- $l cov_x$ is the least probable execution paths coverage ratio;
- $rIndex$ is the adjustment coefficient and is rescaled to avoid its excessive dominance over other factors;
- D_x is the number of executions of danger functions or danger operations;

3 Theory Analyses and Experiments Verification

3.1 Theory Analyses

Let the path depth of the input spot of the binary program to the possible vulnerability position is L , the constraint conditions at very path node in one path are α_i . The path constraint conditions at the every node in one path which are resulted by the input $X_{reg}(x_i)$, the problem of the according test data generation of the GA can turn into the Minimum problem of the function $f(X) = \sum_{i=1}^L |g(x_i) - \alpha_i|$.

For the genetic algorithm technology, we let the probability of the change of the bits in the input $X_{(i)}$ that caused by GA operation is p_m . The probability that the input $X_{(1)}$ is turned into the input $X_{(2)}$ is $p_m^k (1 - p_m)^{n-k}$, where there are k bits differences between the input $X_{(1)}$ and the input $X_{(2)}$. For any input $X_{(i)}$, the probability which the input $X_{(i)}$ is turned into the optimal solution of the minimum problem of $f(X) =$

$\sum_{i=1}^L |g(x_i) - \alpha_i|$ is $p_m^k (1 - p_m)^{L-k}$, where there are k bits differences between the input $X_{(i)}$ and the optimal solution.

For the Fuzzing test technology, we let that there is one input position in the binary program, the path depth of the input spot to the possible vulnerability position is L and the number of nodes in the every layer is m_i , where $1 \leq i \leq L, m_i \geq 1, m = \min_{1 \leq i \leq L} \{m_i\}$. For any input $X_{(i)}$, the probability which the input $X_{(i)}$ is turned into the optimal solution of the minimum problem of the function $f(X) = \sum_{i=1}^L |g(x_i) - \alpha_i|$ is $\frac{1}{\prod_{i=1}^L m_i}$. The maximum of the $\frac{1}{\prod_{i=1}^L m_i}$ is $\frac{1}{m^L}$.

In order to compare the efficiency of GA and Fuzzing technology, we let $y_1 = p_m^k (1 - p_m)^{L-k}, y_2 = \frac{1}{m^L}$, where $k=30, p_m = 0.000001$. From the Figure 3, we can see that the transfer probability of GA is higher than the max transfer probability of Fuzzing technology. The transfer probability is the probability that the input $X_{(i)}$ is turned into the optimal solution of the minimum problem of the function $f(X) = \sum_{i=1}^L |g(x_i) - \alpha_i|$. When the transfer probability is fixed and L is not more than 4500000, from the figure 3, we can see that the efficiency of the GA is higher than the common Fuzzing technology.

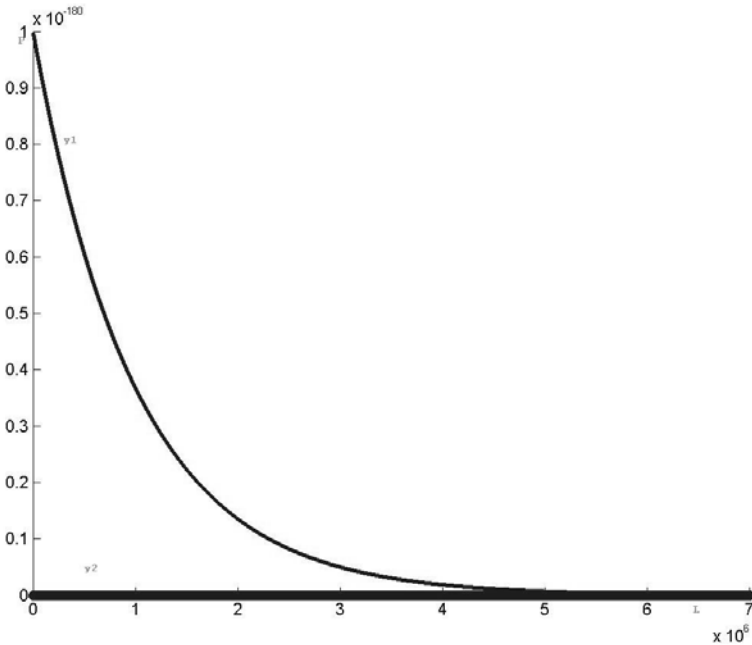


Fig. 3. The transfer probability comparison of the GA and Fuzzing

3.2 Experiments Verification

Our experiments are conducted on a machine with Intel Core 2 Duo CPU at 2.13GHz and 3GB memory, running Windows XP Professional SP3. The debugger module is

build with the windows API. The static analysis module is built on the IDApro, which is a plugin of IDApro.

Table 1. Compare of the Genetic Algorithm and Fuzzing Technology

Method	Initial population size	Iteration number	Run time(s)
Genetic algorithm	200	1057	10678
Fuzzing	200	—	>43200

To verify the effectiveness and efficiency of our system, we test a small binary program which contains integer overflow vulnerability. We compare the iteration number and run time that the genetic algorithm needs and the common Fuzzing technology needs. From the Table I, we can see that the runtime of the method proposed in this paper is 10678s when the integer overflow vulnerability in the binary executables is hit. The common Fuzzing technology has not hit the integer overflow vulnerability in the binary executables in the 43200s. So, the method in this paper is more efficient than the common Fuzzing technology.

4 Conclusions

In this paper, we present a system for detecting the integer overflow in the binary code. The system is combined with debugger module, static analysis module and genetic algorithm module. We use the fitness function of GA to guide the test data generation, at the same time we use the static analysis module to provide the information that the genetic algorithm needs. Contrast with the common fuzzing technique, our system can generate more targeted test data to crash the test program directed by the fitness function. Our system works at the binary level, which does not need the source code, it can greatly help people who do not have the source code to study the software vulnerabilities.

Acknowledgements. We thank the anonymous reviewers for their hard work, helpful comments and suggestions. This research was supported by the National Natural Science Foundation of China under Grant No.61170628 and No.61070207. This research was also supported in part by the National High Technology Research and Development program of China (No. 2007AA01Z466).

References

1. http://cwe.mitre.org/documents/vuln-trends/index.html#overall_trends
2. <http://www.adobe.com/support/security/bulletins/apsb10-02.html>

3. <http://www.debian.org/security/2009/dsa-1912>
4. <http://www.phrack.org/issues.html?issue=60&id=10>
5. Miller, B.P., Fredrikson, L., So, B.: An Empirical Study of the Reliability of UNIX Utilities. *Comm. of the ACM* 33(12), 32 (1990)
6. Miller, B., Koski, D., Lee, C.P., Maganty, V., Murthy, R., Natarajan, A., Steidl, J.: FuzzRevisited: A Reexamination of the Reliability of UNIX Utilities and Services. Technical report, University of Wisconsin-Madison (April 1995)
7. Goldberg, D.: Genetic algorithms in search, optimization and machine learning. Addison-Wesley Publishing Company, Reading (1989)
8. Tracey, N., Clark, J., et al.: Automated test-data generation for exception conditions. *Software-Practice and Experience* 30(1), 61–79 (2000)
9. Del Grosso, C., Antoniol, G., et al.: Detecting buffer overflow via automatic test input data generation. *Computers & Operations Research* 35, 3125–3143 (2008)
10. Sparks, S., Embleton, S., Cunningham, R., et al.: Automated vulnerability analysis: Leveraging control flow for evolutionary input crafting. In: *Proc of the 23rd Annual Computer Security Applications Conference*, pp. 477–486 (2007)
11. IDA Pro. Home page [EB/OL]. [2010-08-01], <http://www.datarescue.com/>
12. Holland, J.: *Adaptation in natural and artificial systems*. University of Michigan Press (1975)

Case-Based Reasoning Genetic Algorithm for Rectangle and Circle Packing Problem with Equilibrium Constraints

Ziqiang Li and Meng-juan Dong

School of Information & Engineering, Xiangtan University, Xiangtan 411105, China
xtulzq@hotmail.com

Abstract. The circle and rectangle packing problem with equilibrium constraints is difficult to solve in polynomial time. In this paper a case based reasoning genetic algorithm is presented to improve its solving efficiency and accuracy. Its main idea is that the initial population of genetic algorithm consists of the reasoning solutions attained from the case-solutions and its non feasible layout solutions generated randomly; in each iteration step of the evolutionary process, a non-isomorphic layout pattern solution of the current optimal individual is constructed automatically and replaces the worse one in the population. The experimental results demonstrate that the proposed algorithm has higher computational speed and accuracy than HCIGA.

Keywords: Circles and rectangles packing problem, Case-based reasoning, Genetic algorithm, Non-isomorphic layout pattern.

1 Introduction

The layout problem, such as the layout scheme design of the satellite module [1] is concerned with placing a set of instruments and equipments into specified space so that the layout scheme satisfies the corresponding constraints, it is called the layout design problem with equilibrium constraints and is surprisingly difficult to solve due to NP nature. Therefore, it has been researched for many years and some algorithms (for example, heuristic method, evolutionary algorithms, graph theory and so on) have been proposed. The evolutionary algorithm (for example, genetic algorithm (short for GA) [2], ant colony optimization [3]) is often used in solving this problem.

There is the deficiency of slower convergence speed and premature phenomenon for GA [4] so that its solving precision is very difficult to be improved. So, scholars probe out many strategies to improve the performance of GA. For example, Cho [5] proposed an interactive genetic algorithm for modeling design and image retrieval in 2002. Its idea is that fitness value of individual is evaluated by human with creative thinking at each iterative step of evolution; in order to combine it with human intelligence, Qian et al[6] proposed a human computer interactive genetic algorithm (short for HCIGA) and applied to layout problem with performance constraints in 2001. Its idea is that the artificial design schemes are constantly coded into unified normalization string, and are added timely into the population of GA by replacing the individual with worse fitness

in the process of iteration. Thus, both human and computer can exert their respective advantages to the utmost; in 2007 Bo Jin [7] combined the case-based retrieval and evolutionary algorithm to solve layout problem of spacecraft modules, which improves performance of the solution. In this paper, using non-isomorphic layout pattern (short for NLP [8]), a case-based reasoning (short for CBR) GA is presented to solve the circle and rectangle packing problem with equilibrium constraints. Its idea is that through layout CBR, its reasoning solutions are obtained from the case database and taken as a part of initial population individuals so that initial population includes elite solutions; the rest part generated randomly. In addition, in each step of iteration, the population individual with worse fitness is replaced by the layout pattern solution of the current optimal solution (the optimal individual). The proposed algorithm can better avoid falling into the local optimal in the solving process of GA.

For the layout pattern, Li [9] gave out isomorphic and non-isomorphic definition, and described their constructing and judging approaches in 2003. But it more adapts to the dense case of object distribution. Teng & Li [8] presented quasi-boundary line method (short for QBLA) to construct NLP solution in 2006. It adapts to both dense and sparse situation of the object distribution.

This paper is organized as follows. The reasoning rule based on the layout case is presented in Section 2. The case-based reasoning layout GA is proposed in Section 3. The experiment results are in section 4. Section 5 is summary of this paper.

2 Layout Case Based Reasoning

CBR [10] possesses many advantages such as fast reasoning, remembering, easy explanation etc. It is widely applied in a lot of fields. For example, Bo Jin [7] combined CBR and evolutionary algorithm to solve the layout design of spacecraft modules for improving performance of the solution in 2007; Park [11] presented a interactive CBR method considering proximity from the cut-off point for medical diagnosis in 2007. According to application viewpoint, CBR system can be classified into two types: explanation and resolution. The difference is that for the former the reasoning solution of the problem to be solved is directly obtained from all case solutions by comprising and judging; for the latter it is obtained by newly constructing based on the case solution amended or modified. In the proposed CBR, for all of the case solutions, if its model is the same as the layout problem, then it is directly taken as its reasoning solution, if approximate, then using them to anew generate its reasoning solution. CBR in the paper belongs to the hybrid type.

2.1 The Layout Case Storage

Taking the layout problem of satellite module (see section 4) as example, the case model¹ is described. All layout cases are stored in the database with the record form.

¹ The layout model of case: type of constraint (char), non-equilibrium and radius of optimal(double), The shape of container (char), The number and size of circle and rectangle (int and double).

2.2 The Proposed Layout CBR

(a) Related definition. To describe CBR of 2-D layout problem in this paper, we give several related definitions at first.

Definition 1. Let D_C denote the distance between the containers of the layout problem and the case, and their shape is the same then $D_C=0$; otherwise $D_C=1$.

Definition 2. Let N_{R1} and N_{R2} denote the number of the rectangles of the layout problem and that of the case respectively, N_{C1} and N_{C2} denote its number of the circles and that of the cases respectively, then the distance D_{Obj} between two sets of objects can be calculated by formula (1). Where, for $i=1,2,\dots, N_{C1}$, if its i th circle of the layout problem can be found in the case, then $a_i=0$; otherwise $a_i=1$; for $j=1,2,\dots, N_{R1}$, if its j th rectangle can be found in the cases, then $b_j=0$, otherwise $b_j=1$.

$$D_{Obj} = \frac{1}{2} \left(\frac{|N_{C2} - N_{C1}| + |N_{R2} - N_{R1}|}{\max(N_{C2}, N_{C1}) + \max(N_{R1}, N_{R2})} + \frac{\sum_{i=1}^{N_{C1}} a_i + \sum_{j=1}^{N_{R1}} b_j}{\max(N_{C1}, N_{C2}) + \max(N_{R1}, N_{R2})} \right) \quad (1)$$

Definition 3. Let I_p and I_c denote the performance indices of the layout problem and the case respectively, then the distance D_J between their performance indices can be calculated by formula (2).

$$D_J = \frac{|I_p - I_c|}{\max(I_p, I_c)} \quad (2)$$

Definition 4. Let D denote the distance between the layout problem and case, then D can be calculated by formula (3). Where, W_C , W_{Obj} and W_I are the weighted factor of D_C , D_{Obj} and D_J respectively and $W_C + W_{Obj} + W_I = 1$.

$$D = W_C D_C + W_{Obj} D_{Obj} + W_I D_J \quad (3)$$

(b) Case retrieval and reasoning

The distance D ($0 \leq D \leq 1$) of each case is calculated according to formula (3). Taking less threshold ε ($\varepsilon > 0$), then the case that $D \leq \varepsilon$ is called the similar cases. Let V_p and V_c denote the size vectors of containers of the layout problem and the case. For the similar case, its reasoning principle is as follows.

(i) If $D = 0$, then takes its layout scheme as reasoning solution of the layout problem.

(ii) If $0 < D \leq \varepsilon$ and $D_C = 0$, $V_p = V_c$, but $N_{C1} \neq N_{C2}$ or $N_{R1} \neq N_{R2}$, then the reasoning solution is generated by matching the case solution. That is, for the object with the same size, its mass center is that of the corresponding object in the case solution; for the object with the same shape, its mass center is that of the object with approximate size in the case solution; For the other objects, then their mass centers are generated randomly.

(iii) If $0 < D \leq \varepsilon$, $D_C = 0$, $N_{C1} = N_{C2}$, $N_{R1} = N_{R2}$ but $V_p \neq V_c$, let k is the scale coefficient of their container size, X_c is the case solution, then the reasoning solution $X_p = k X_c$.

(iv) If $0 < D \leq \varepsilon$, $D_C = 0$ but $V_p \neq V_c$, $N_{C1} \neq N_{C2}$, $N_{R1} \neq N_{R2}$, then the reasoning solution of the layout problem generates by executing (ii) and (iii) in turn.

(v) If none of the 4 cases above is satisfied, then by executing (ii) to obtain X_c . The reasoning solution X_p is the optimal solution of the proposed algorithm for X_c .

3 The Case-Based Reasoning Layout Genetic Algorithm

Theoretically, GA with optimal maintaining strategy is global convergent. But it easily appears premature phenomenon in practical. So, scholars have been looking for better evolutionary strategy for avoiding the phenomenon. Research results show that performance of the solution is closely related to the initial population and the population diversity in process of iteration. By far the former has strategies as follows:

- i) Directly constructs through human-computer interaction [6].
- ii) Uniformly distributed construction [12].
- iii) It consists of random initial solutions and their NLP solutions [13].

In this paper, QBLA is used in constructing of NLP based on the introduction in Section 1. Let population size be M , Number of reasoning resolutions be m ($0 < m \ll M$), then the initial population is generated according to following method:

Firstly, $[(M-m)/m_1]$ NLP solutions are constructed by QBLA for each reasoning resolution, where $[(M-m)/m_1]$ is the integer part of $(M-m)/m_1$ and m_1 is an integer with $m < m_1 < 3m/2$. Afterward, $M-m-[(M-m)/m_1] \times m$ random solutions are generated. The initial population of GA in the proposed algorithm consists of the reasoning solutions and their NLP solutions and random solutions. Because the elite solution is involved, its quantity is improved.

To enhance the population diversity, NLP solution of the best individual is constructed through QBLA and the worse individual is replaced in each iteration step of GA. The method is as following: let the number of objects be n . the object A_i and A_j are generated randomly ($1 \leq i, j \leq n$ and $j \neq i$). Isomorphic control zone (see [8]) $D_f(A_i)$ and $D_f(A_j)$ of objects A_i and A_j is constructed by QBLA respectively (see Figure 1). If the mass center C_j of object A_j is not in the control zone $D_f(A_i)$ of A_i , then A_i is moved to make C_i in $D_f(A_j)$, at the same time A_j is moved to make C_j in $D_f(A_i)$.

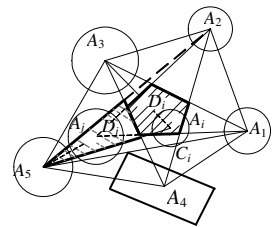


Fig. 1. Non-isomorphic control zone of objects

The main steps of the proposed algorithm are as following:

- Step1. Constructing initial population of GA through CBR and NPL construction;
- Step2. Crossover, mutation and selection operation are executed on the population in turn. The optimal individual X_{opt} of the population is calculated.
- Step3. If X_{opt} satisfies requirement or GA has implemented maximum iterative times, then go to Step4, otherwise iterative times adds 1, and the worse one in the population is replaced by NPL solution of the current optimal individual, go to Step2;
- Step4. Output the optimal solution, algorithm ends.

4 The Experiment Results

Our algorithm run on PC with 933M CPU/512M Memory, run environment of GA and HCICA is computer with 566M CPU/128M Memory.

Problem: its background is the satellite module design. Try to place 5 rectangles: 1(length=365, wide =365, mass= 4.0), 2(425,300,4.5), 3(250,225,3.6), 4(380,280,3.9), 5(250,175,1.8) and 5circles: 6(225, 4.0),7(200,2.0), 8(170,2.6), 9(75,1.4), 10(200,2.5) (the circle 10 can't move) in the circular container with radius $R_c=750\text{mm}$ in term of following conditions: 1) there is non-overlapping between them and between them and container; 2) the static non-equilibrium should be smaller than ϵ_j . 3) The radius of envelope circle is the smallest. Its mathematical model is showed in formula (4). Where, $J(X)$ is the static non-equilibrium of X , $R(X)$ is the smallest radius of envelope circle, A_i and A_j are circle or rectangle respectively.

$$\text{Find } X=\{(x_i,y_i) \quad i=1,2,\dots,n\}$$

$$J(X) = \sqrt{\left(\sum_{i=1}^n m_i x_i\right)^2 + \left(\sum_{i=1}^n m_i y_i\right)^2} \leq \delta \tag{4}$$

$$R(X)=\min\{(x_i^2 + y_i^2)^{1/2} + d_i\} \text{ s t. } \text{int } A_i \cap A_j = \emptyset \quad i,j=1,2,\dots,n \text{ and } i \neq j$$

Solving: According to its layout model: take $\epsilon=0.15$, $W_j=0.4$, $W_c=0.25$, $W_{obj}=0.35$, the static non-equilibrium: $\epsilon_j=10g.mm$; the offset of mass center: $\epsilon_m=1.2mm$; container type: circle. Using CBR in section 2, we obtained 6 reasoning solutions $X_1(R=748.04, \epsilon_j=0.946, \epsilon_m=0.782)$, $X_2(R=747.89mm, \epsilon_j=0.179, \epsilon_m=0.428)$, $X_3(R=745.46 mm, \epsilon_j=0.003, \epsilon_m=0.006)$, $X_4(R=740.200mm, \epsilon_j=2.020, \epsilon_m=4.830)$, $X_5(R=746.67mm, \epsilon_j=0.730, \epsilon_m=1.470)$, $X_6(R=748.81mm, \epsilon_j=1.790, \epsilon_m=4.290)$. The layout schemes of the reasoning solutions X_1-X_6 are shown in the figure2 (a)~(f).

The parameters of GA are as follows: roulette selection, two points cross and basic bit mutation. The probability of crossover and mutation is 0.99 and 0.30 respectively, the size of population is 60 and the maximal number of iterations is 10000. By using

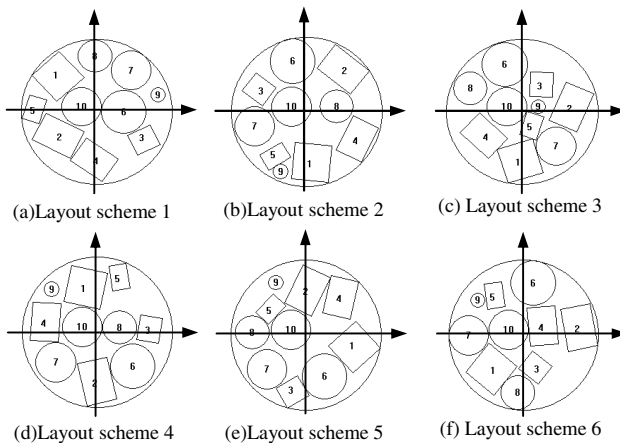


Fig. 2. 6 parameter layout schemes for the example

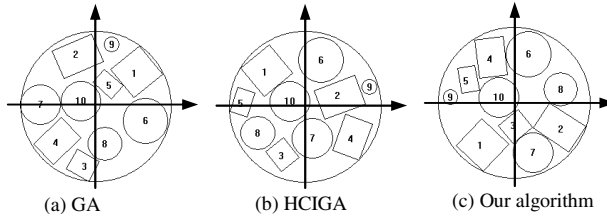


Fig. 3. The layout schemes of three algorithms for the example (a) GA, (b) HCIGA, (c) Our algorithm

QBLA, we structure the 30 NLP solutions of 6 reasoning solutions and 24 random solutions to consist of the initial population. We use penalty-function method to solve the solution 50 times. The optimal layout schemes of three algorithms are shown in Figure 3 (a)~(c). The coordinates of mass center of the optimal layout schemes of GA, HCIGA and the proposed algorithm are shown in table 1, table 2 is the performance comparison of their optimal solutions. It is known from table 2 that the solving quality and efficiency of the proposed algorithm are superior to that of GA and HCIGA.

Table 1. The mass-center coordinates of optimal solutions of three algorithms for the example

Objects num	GA			HCIGA			Our algorithm		
	x_i	y_i	a_i	x_i	y_i	a_i	x_i	y_i	a_i
1	403.682	343.213	0.8877	-376.7859	358.8357	0.8863	-313.60	-414.7099	2.3287
2	-183.84	495.810	2.7325	339.4017	93.27236	2.7830	442.481	-244.5039	0.5883
3	-130.08	-585.97	0.5161	-223.538	-472.449	0.8856	12.9763	-232.9016	2.4577
4	-386.12	-342.34	2.3779	462.8917	-299.711	1.9191	-224.61	445.7915	1.4520
5	117.749	222.100	0.7342	-615.713	46.08496	1.87334	-450.15	229.5832	1.3940
6	495.647	-147.399		157.3876	460.683		145.9251	475.1838	
7	-549.85	12.423		73.8161	-312.03		186.873	-494.980	
8	92.7380	-373.798		-461.940	-262.225		458.794	130.003	
9	157.369	613.510		633.165	192.60		-623.259	51.458	
10	-150.00	50.000		-150.00	50.000		-150.000	50.000	

Table 2. The performance comparison of three algorithms of the example

method	R/mm	$\varepsilon_r/g.mm$	ε_w/mm	δ_r/mm	δ_w/mm	t/s
GA	749.99	0.2953	0.7071	-0.1789	0.2124	313.0
HCIGA	746.36	0.0043	0.0104	-0.0024	-0.0041	217.5(Human40)
Our algorithm	741.09	0.0039	0.0094	-0.0010	0.0040	180.3

5 The Conclusion

In this paper, by combining GA with CBR, an algorithm of circle and rectangle packing problem with equilibrium constraints is suggested. In this algorithm, reasoning solutions obtained through CBR and their NLP solutions are involved into initial population to improve its quantity and at each iteration step, by constructing NLP solutions

of optimal individual and using it to replace the individual with worse fitness the diversity of population is increased to better jump out of the local optimal solution in the situation of no manual intervention. Therefore its solving quality and efficiency and automation degree are improved compared with HCIGA,. Numerical experiment results indicate that the performance of the proposed algorithm is superior to GA and HCIGA.

Acknowledgments. This work is supported by the Hunan Provincial Natural Science Foundation of China (Grant No. 11JJ6050) and Research Foundation of Education Bureau of Hunan Province, China (Grant No. 11A120) and the Doctor Startup Foundation of Xiangtan University of China (Grant No. 09QDZ18).

References

1. Pierre, M.G., Georges, M.F.: A GA based configuration design optimization Method. *J. of Mechanical Design* 26, 6–14 (2004)
2. Liu, Z.W., Teng, H.F.: Human-computer cooperative layout design method and its application. *Computers & Industrial Engineering* 55, 735–757 (2008)
3. Sun, Z.G., Teng, H.F.: Optimization layout solution of satellite module. *J. of Engineering Optimization* 35, 513–529 (2003)
4. Rudolph, G.: Convergence analysis of canonical genetic algorithms. *IEEE Transactions on Neural Networks* 5, 96–100 (1994)
5. Cho, S.B.: Towards creative evolutionary systems with interactive genetic algorithm. *Application Intelligent* 16, 129–138 (2002)
6. Qian, Z.Q., Teng, H.F., Sun, Z.G.: Human-computer interactive genetic algorithm and its application to constrained layout optimization. *Chinese J. of Computers* 24, 553–559 (2001)
7. Jin, B., Teng, H.F.: Case-based evolutionary design approach for satellite module layout. *J. of Scientific and Industrial Research* 66, 989–994 (2007)
8. Teng, H.F., Li, Z.Q., Shi, Y.J., et al.: An approach to constructing isomorphic or non-isomorphic layout pattern. *Chinese J. of Computers* 29, 987–991 (2006)
9. Li, G.Q., Teng, H.F.: Isomorphic and Non-Isomorphic layout patterns of packing problems. *Chinese J. of Computers* 26, 1248–1254 (2003)
10. Earltta, B.: An introduction to Case Based Reasoning. *AI. Expert*, 43–49 (1991)
11. Park, Y.J., Kim, B.C.: An interactive case-based reasoning method considering proximity from the cut-off point. *Expert Systems with Application* 33, 903–915 (2007)
12. Wang, Y.S., Shi, Y.J., Teng, H.-F.: An improved scatter search for circles packing problem with the equilibrium constraint. *Chinese J. of Computers* 32, 1214–1221 (2009)
13. Li, G.Q.: Research on the theory and methods of layout design and its applications (Phd dissertation), Dalian University of Technology (2003)

Feature Discriminability for Pattern Classification Based on Neural Incremental Attribute Learning

Ting Wang^{1,2}, Sheng-Uei Guan², and Fei Liu³

¹ Department of Computer Science, University of Liverpool, Liverpool L69 3BX, UK

² Department of Computer Science and Software Engineering,
Xi'an Jiaotong-Liverpool University, Suzhou 215123, China

³ Department of Computer Science & Computer Engineering, La Trobe University,
Victoria 3086, Australia

ting.wang@liverpool.ac.uk, steven.guan@xjtlu.edu.cn,
f.liu@latrobe.edu.au

Abstract. Feature ordering is important in Incremental Attribute Learning where features are gradually trained in one or more size. Apart from time-consuming contribution-based feature ordering methods, feature ordering also can be derived by filter criteria. In this paper, a novel criterion based on a new metric called Discriminability is presented to give ranks for feature ordering. Final results show that the new metric not only is applicable for IAL, but also exhibits better performance in lower error rates.

Keywords: neural networks, feature ordering, incremental attribute learning.

1 Introduction

In pattern classification, the number of features (attributes) indicates the complexity of a problem. The more features in a problem, the more complex it is. To solve complex classification problems, some dimensional reduction strategies like feature selection have been employed [1, 2]. However, these methods are invalid when the feature number is huge and almost all features are crucial simultaneously. Thus feature reduction is not the ultimate technique to cope with high dimensional problems.

A strategy for solving high-dimensional problems is “divide-and-conquer”, where a complex problem is firstly separated into smaller modules by features and integrated after each module is tackled independently. Incremental Attribute Learning (IAL) is an example of that. It is applicable for solving classification problems in machine learning [3-6]. Previous studies show that IAL based on neural networks obtains better results than conventional methods [3, 7]. For example, in Guan’s studies, compared with traditional methods [5, 6], classification error rates of Diabetes, Thyroid and Glass, three machine learning datasets from University of California, Irvine (UCI), derived by neural IAL were reduced by 8.2%, 14.6% and 12.6%, respectively [8].

However, because IAL incrementally imports features into systems, it is necessary to know which feature should be introduced in an earlier step. Thus feature ordering becomes a new preprocess apart from conventional preprocess like feature reduction.

Previous studies of neural IAL presented contribution-based feature ordering method, where feature ordering was derived after each feature is solely employed to classify all outputs by neural networks. The result of each denotes every feature's ability for discrimination. However, such a wrapper is more time-consuming than filter [9]. Thus it is necessary to study on feature ordering based on filter methods. mRMR[10], a filter feature selection criterion, is successfully employed to get optimal feature ordering, although not all performance of this approach is better than that obtained by wrapper methods [11]. Therefore, whether there are some metrics existing, which can effectively rank features and produce accurate classification results, is worthy of studying.

In this paper, a new statistical metric called Discriminability is presented for feature ordering. It is defined by pattern distribution and will be checked for applicability and accuracy by ITID, a neural IAL algorithm. In Section 2, ITID will be reviewed and Discriminability will be presented in Section 3; Benchmarks will be analyzed in Section 4; conclusions will be drawn in section 5 with outlines of future works.

2 IAL Base on Neural Networks

In previous studies, based on some predictive methods like neural networks, IAL is proved that it is feasible to solve multi-dimensional classification problems. ITID [12], a neural IAL based on ILIA1 and ILIA2 [7], is shown applicable for classification. It divides all input dimensions into several sub-dimensions, each of which corresponds to an input feature. Instead of learning input features altogether as an input vector in training, ITID learns inputs through their corresponding sub-networks one after another and the structure of neural networks gradually grows with an increasing input dimension as shown in Figure 1. During training, information obtained by a new sub-network is merged together with the information obtained by the old network. Such architecture is based on ILIA1. After training, if outputs are collapsed with an additional network sitting on the top, links to the collapsed output units and all input units are built to collect more information from inputs. Such a reformed network is based on ILIA2 and is shown in Figure 2. Finally, a pruning technique is adopted to find the appropriate network architecture. With less internal interference among input features, ITID achieves higher generalization accuracy than conventional methods [12].

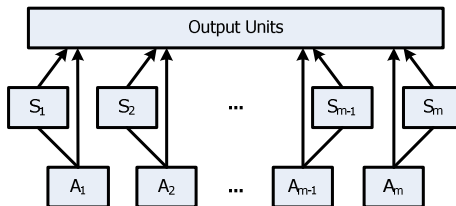


Fig. 1. The basic network structure of ITID

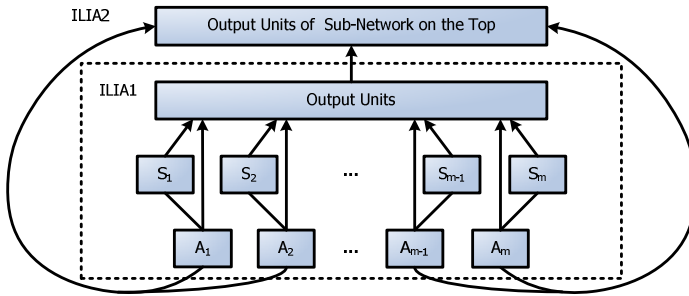


Fig. 2. The network structure of ITID with a sub-network on the top

3 Discriminability for Feature Ordering

Feature ordering is seldom used in conventional methods where features are trained in one batch. In previous studies, feature ordering was derived by single feature’s contribution. However, it is time-consuming. To cover such a shortage, filter feature ranking approaches are employed to arrange feature ordering based on discrimination ability. The greater discrimination ability a feature has, the more prior this feature should be trained. Although filter feature ordering method is similar to filter feature selection, and can be simultaneously computed with feature reduction, there are still one key difference between feature ordering and feature reduction: feature reduction usually scraps useless features or reduces the weights of useless features, while feature ordering does nothing but give a sequence of features by discrimination ability.

In classification problems, classes can be separated by input features, because contribution of different inputs in classification is variable. Some features are good at distinguishing one class from others, while others are weak in this categorization. Thus different features have different discrimination abilities in classification. Such ability can be regarded as a rank for feature ordering.

Definition 1. Discriminability refers to the discriminating capacity of one input feature x_i in distinguishing all output features $\omega_1, \omega_2, \dots, \omega_n$, where x_i is the i^{th} feature in the set of inputs, n is the number of output features.

Therefore, two kinds of metrics can be employed to compute Discriminability. One is the standard deviation of sample means from all output features in one input dimension, and the other is the sum of standard deviation of each output feature in one input dimension. Obviously, the former indicates the scatterance of all samples classified by classes in one input dimension, where the more symmetrical of the scatterance is, the less availability of this feature is. Moreover, the latter denotes the integrative data gathering level of each output in one input dimension. Manifestly, the more dispersive of data gathering is on one dimension, the more difficult of this feature for this classification is. Specifically, let $\mathbf{X} = [x_1, \dots, x_m]$ the pool of input and $\mathbf{\Omega} = [\omega_1, \dots, \omega_n]$ the pool of output, where $x_i (1 \leq i \leq m)$ is the i^{th} input features in \mathbf{X} and $\omega_j (1 \leq j \leq n)$ is the j^{th} output feature in $\mathbf{\Omega}$, Discriminability can be calculated by

$$D(x_i) = \frac{std[\mu_1(x_i), \dots, \mu_n(x_i)]}{\sum_{i=1, j=1}^{i=m, j=n} std_j(x_i)} \tag{1}$$

where $\mu_j(x_i)$ is the mean of feature i in output j , $std_j(x_i)$ is its standard deviation, m is the number of input, and n is the number of output. Discriminability provides an indicative feature ordering ranking in two or more output categorization problems.

Feature ordering is a unique and indispensable data preparation job of IAL. When all features are ranked by Discriminability, they should be placed in a descending order, which have been proved that such ordering is more likely to obtain better results [13]. After that, patterns are randomly divided into three different datasets: training, validation and testing [14]. All vectors in these three datasets should be sorted according to the feature ordering and employed for classification by machine learning later.

4 Experiments and Analysis

The proposed ordered IAL using Discriminability was tested on three classification benchmarks from UCI machine learning datasets: Diabetes, Cancer, and Glass where the brief information of these three datasets has been shown in Table 1.

Table 1. Brief information of benchmarks

	Diabetes	Cancer	Glass
Pattern Number	768	699	214
Input Number	8	9	9
Output Number	2	2	6

All patterns of these three benchmarks were randomly divided into three sets: training (50%), validation (25%) and testing (25%), where training data were firstly used to rank feature ordering based on Discriminability, and ITID-D, ITID based on Discriminability for feature ordering, was employed for classification. The performance of the utilization of Discriminability is evaluated based on the comparison of error rate with other three approaches: mRMR (Difference), mRMR (Quotient), wrappers and conventional batch training method. Table 2, 3, and 4 shows the Discriminability and input index calculated by equation (1) of Diabetes, Cancer, and Glass, respectively, and corresponding results are demonstrated with comparison with other classification methods in Table 5, 6, and 7.

Table 2. Discriminability of Diabetes

Ordering	1	2	3	4	5	6	7	8
Discri.	0.3694	0.2630	0.1844	0.1567	0.1437	0.1047	0.0960	0.0509
Index	2	6	8	7	1	4	5	3

Table 3. Discriminability of Cancer

Ordering	1	2	3	4	5	6	7	8	9
Discri.	0.9888	0.9566	0.8725	0.7793	0.7039	0.6895	0.6604	0.6260	0.3534
Index	3	2	6	7	1	8	4	5	9

Table 4. Discriminability of Glass

Ordering	1	2	3	4	5	6	7	8	9
Discri.	0.3226	0.2605	0.1716	0.1566	0.1514	0.0976	0.0802	0.0764	0.0542
Index	3	8	4	2	6	5	9	1	7

Table 5. Result comparison of Diabetes

	Feature Ordering	Error Rate (ILIA1)	Error Rate (ILIA2)
ITID-D	2-6-8-7-1-4-5-3	21.84896%	22.39583%
mRMR-Difference	2-6-1-7-3-8-4-5	22.86459%	23.5677%
mRMR-Quotient	2-6-1-7-3-8-5-4	22.96876%	23.82813%
wrappers	2-6-1-7-3-8-5-4	22.96876%	23.82813%
Conventional method		By batch: 23.93229%	

Table 6. Result comparison of Cancer

	Feature Ordering	Error Rate (ILIA1)	Error Rate (ILIA2)
ITID-D	3-2-6-7-1-8-4-5-9	1.695405%	1.72414%
mRMR-Difference	2-6-1-7-3-8-5-4-9	2.29885%	1.580462%
mRMR-Quotient	2-6-1-7-8-3-5-4-9	2.29885%	1.810347%
wrappers	2-3-5-8-6-7-4-1-9	2.4999985%	1.925289%
Conventional method		By batch: 1.867818%	

Table 7. Result comparison of Glass

	Feature Ordering	Error Rate (ILIA1)	Error Rate (ILIA2)
ITID-D	3-8-4-2-6-5-9-1-7	34.81133%	28.96228%
mRMR-Difference	3-2-4-5-7-9-8-6-1	39.05663%	35.09436%
mRMR-Quotient	3-5-2-8-9-4-7-6-1	35.28304%	31.50946%
wrappers	4-2-8-3-6-9-1-7-5	36.4151%	33.11322%
Conventional method		By batch: 41.226405%	

According to the results, ITID-D exhibits the best performance with lowest classification error rate in Diabetes (ILIA1, ILIA2), Cancer (ILIA1) and Glass (ILIA1, ILIA2), and also obtained the secondary lowest error rate in Cancer (ILIA2).

5 Future Work and Conclusions

IAL is a novel approach which trains input attributes gradually in one or more sizes. Feature ordering in training is unique to IAL. In previous studies, IAL feature

ordering was derived by wrapper methods which are more time-consuming than filters like Discriminability and mRMR. Moreover, experimental results also demonstrated that the feature ordering obtained by Discriminability can exhibit better performance than other methods. Furthermore, some classification results derived by ITID using Discriminability are better than those produced by mRMR, a wide used filter criterion. Nonetheless, there are a number of further studies needed to be done in the future. For example, although Discriminability can attain better performance than the other approaches, how to obtain feature ordering with an optimal pattern recognition result is still unknown. Generally, using Discriminability in feature ordering is applicable for time saving and classification rate enhancing in neural IAL-based pattern classification problems.

Acknowledgments. This research is supported by National Natural Science Foundation of China under Grant 61070085.

References

1. Liu, H.: Evolving feature selection. *IEEE Intelligent Systems* 20(6), 64–76 (2005)
2. Weiss, S.H., Indurkha, N.: *Predictive data mining: a practical guide*. Morgan Kaufmann Publishers, CA (1998)
3. Chao, S., Wong, F.: An incremental decision tree learning methodology regarding attributes in medical data mining. In: *Proc. of the 8th Int'l Conf. on Machine Learning and Cybernetics*, Baoding, pp. 1694–1699 (2009)
4. Agrawal, R.K., Bala, R.: Incremental Bayesian classification for multivariate normal distribution data. *Pattern Recognition Letters* 29(13), 1873–1876 (2008)
5. Guan, S.U., Liu, J.: Feature selection for modular networks based on incremental training. *Journal of Intelligent Systems* 14(4), 353–383 (2005)
6. Zhu, F., Guan, S.U.: Ordered incremental training for GA-based classifiers. *Pattern Recognition Letters* 26(14), 2135–2151 (2005)
7. Guan, S.U., Li, S.: Incremental learning with respect to new incoming input attributes. *Neural Processing Letters* 14(3), 241–260 (2001)
8. Guan, S.U., Li, S.: Parallel growing and training of neural networks using output parallelism. *IEEE Trans. on Neural Networks* 13(3), 542–550 (2002)
9. Bermejo, P., de la Ossa, L., Gámez, J.A., Puerta, J.M.: Fast wrapper feature subset selection in high-dimensional datasets by means of filter re-ranking. *Knowledge-Based Systems* (accepted, 2011)
10. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(8), 1226–1238 (2005)
11. Wang, T., Guan, S.U., Liu, F.: Ordered Incremental Attribute Learning based on mRMR and Neural Networks. *International Journal of Design, Analysis and Tools for Integrated Circuits and Systems* 2, 86–90 (2011)
12. Guan, S.U., Liu, J.: Incremental neural network training with an increasing input dimension. *Journal of Intelligent Systems* 13(1), 43–69 (2004)
13. Guan, S.U., Liu, J.: Incremental Ordered Neural Network Training. *Journal of Intelligent Systems* 12(3), 137–172 (2002)
14. Ripley, B.D.: *Pattern Recognition and Neural Networks*. Cambridge University Press, UK (1996)

Selective Ensemble Approach for Classification of Datasets with Incomplete Values

Yan Wang¹, Yi Gao¹, Ruimin Shen¹, and Fan Yang²

¹ School of Continuing Education, Shanghai Jiao Tong University, Shanghai 200240, China

² Chongqing Academy of Science & Technology, Chongqing 401123, China

wangyan8383@sjtu.edu.cn

Abstract. In some research situations, we often have to classify data with incomplete values which affect the learning performance of classifiers. Although various classification algorithms have been proposed, most of them are short of the ability to deal with incomplete data. This paper proposes a novel approach based on selective ensemble for classifying incomplete data. The method finds the local complete patterns for which the feature values are complete and trains multiple component learners for each local complete subset. Then, it combines the outputs of the classifiers. The method needs no assumption about the incomplete mechanism that is necessary for previous methods. The proposed method is evaluated by three datasets from the UCI Machine Learning Repository. The experiments results show that classification accuracy of the proposed method is superior to those of widely used imputations and deletion method.

Keywords: Incomplete data, Local complete pattern, Selective ensemble classification.

1 Introduction

The incomplete data are common occurrences in most applied research, that is, some feature values or the class labels are missing in the process of collecting data. Incomplete data could be caused by various reasons, such as error of data measure, error of data understanding, erroneous human imputation, or restriction of data collecting. Most classification algorithms are short of the ability to approach the incomplete data problem except Naïve Bayes [1] and C4.5 [2]. Naïve Bayes classifier imputes the feature value with special value while C4.5 ignores the instances with missing entries. Both methods may incur potentially dangerous biases in the estimates. Therefore, the reasonable method is deserved discussing in classification problem for incomplete data.

In recent years, many solutions for dealing with incomplete data have been proposed. Qu et al. [3] deleted the instances with incomplete data before constructing the classification model. Although deleting the incomplete data improved the classification performance in some degree, deletion may discard some important information within the incomplete data, especially under the condition of insufficient data. In this paper, we construct the classification model based on the incomplete data.

Boujelben et al.[4] predicted the incomplete data based on evidence theory. However, the evidence function should be learned in advance. Salama et al.[5] constructed the complete dataset by statistics method. Huang et al.[6] filled the missing data by Bayesian model. Nevertheless, Bayesian model needs to know probability distribution, which is difficult to acquire in fact. In some applications, expert experience could be used to form the complete data. However, the method of forecast incomplete data by experts is subjective.

Williams et al.[7] proposed a method for classifying incomplete data with logistic regression. Although the method needed no imputation for the incomplete data, it relied on the assumption of missing at random (*MAR*). However, no way could verify whether the data was *MAR* or not.

Ramoni et al.[8] introduced the Robust Bayes classifier (*RBC*) that made use of intervals bounding all the estimates of related probabilities and ignored the assumption about *MAR*. Although *RBC* method can avoid the biases in some extent, it needed the independent assumption between the attributes and the class. When the assumption was not established, the classification performance would degrade.

To overcome the biases from deletion and imputation method, the proposed approach focuses on constructing classifier on the incomplete data sets directly. Firstly, the proposed method finds the local complete pattern within incomplete data set, and then classifies the incomplete data by an ensemble selection classification on local complete data sets. While merging multi classification results, the method utilizes the complete rate of each data subset and the performance of each individual classifier to replace the usual voting policy.

2 The Proposed Selective Ensemble Approach

2.1 Finding Local Complete Pattern (LCP) from Incomplete Data

This section discusses the learning approach of the local complete pattern within the incomplete data. At first, finding all possible complete feature sets (*CFS*), and then mapping *CFS* into the corresponding complete subsets of samples for which the features values are complete.

Definition 1. A 4-tuple $S = \langle U, F, Y, g \rangle$ is called an incomplete decision system (*IDS*), where U is a non-empty finite set of objects called universe, F is a non-empty finite set of condition features, and Y is a non-empty finite set of decision features where $F \cap Y = \emptyset$. For $\forall f \in F, x \in U, g(x, f)$ denote the value that x holds on feature f . Then, there exists at least $\exists x \in U$ and $\exists f \in F$ that satisfy $g(x, f) = *$. We assume here that the incomplete feature values are denoted by “*”.

Definition 2. Let $IFS = \{f_i\}$ be the incomplete feature set, that is, there exists an object in U whose feature value about IFS is missing while about $F - IFS$ is complete. The constructing process of incomplete feature set is formalized as follows:

$$\exists x_k \in U, \forall f_i \in IFS \wedge \forall f_j \in (F - IFS), \text{ satisfy } g(f_i, x_k) = * \wedge g(f_j, x_k) \neq *(i \neq j).$$

Then $CFS = F - IFS$ is the complete feature set of S . It is easy to deduce that $|CFS| + |IFS| = |F|$, where $|\bullet|$ denotes the number of objects in set.

Example 1. An *IDS* with five condition features and five objects is shown in Table 1. The *IFS* are $IFS1 = \{f_3\}$, $IFS2 = \{f_1, f_3, f_4, f_5\}$, $IFS3 = \{f_4, f_5\}$, $IFS4 = \{f_3\}$. Then the corresponding complete feature sets are

$$CFS1 = S - IFS1 = \{f_1, f_2, f_4, f_5\}, CFS2 = \{f_2\}, CFS3 = \{f_1, f_2, f_3\}, CFS4 = \{f_1, f_2, f_3, f_4\}$$

Table 1. The incomplete decision system

U	f1	f2	f3	f4	f5	y
x1	1	0	*	0	1	1
x2	*	1	*	*	*	0
x3	0	1	1	*	*	0
x4	1	1	*	1	0	1
x5	1	1	1	1	*	1

Definition 3. A 4-tuple $LCP = \langle U', CFS, Y, g \rangle$ is called a local complete pattern, where $CFS \subset F$ is a non-empty finite set of complete feature, $U' \subset U$ is the mapping of CFS into the set of all subsets of the set U , and Y is a non-empty finite set of decision features. For $\forall x_j \in U' \wedge \forall f_i \in CFS$, satisfies $g(f_i, x_j) \neq *$.

Example 2. For the complete feature sets $\{CFS_i\}_{i=1}^4$ shown in Example 1, four local complete patterns $\{LCP_i\}_{i=1}^4$ shown in Table 2 can be found. The example shows that *LCP* can be found from the incomplete dataset easily. Thus, it can be adopted as a pretreated method in classification model constructing.

Table 2. The local complete patterns about Example 1

$LCP1$	f1	f2	f4	f5	y	$LCP2$	f2	y	$LCP3$	f1	f2	f3	y	$LCP4$	f1	f2	f3	f4	y
x1	1	0	0	1	1	x1	0	1	x3	0	1	1	0	x5	1	1	1	1	1
x4	1	1	1	0	1	x2	1	0	x5	1	1	1	1						
						x3	1	0											
						x4	1	1											
						x5	1	1											

(a)

(b)

(c)

(d)

2.2 Selective Ensemble Approach (SECID)

This section discusses the constructing approach of classifier under incomplete data. Considering each *LCP* reflects partial data characteristic from different local complete data views, we construct multiply different classifiers on each *LCP*, and then combine the multiply classification results to predict the class label. The purpose of using ensemble is to integrate different classification decisions in order to get a more reliable

performance. Ensemble methods comprise two main phases, which are the production of the different models and the combination of the models [9]. In general, bagging[10], boosting[11], and stacking[12] are the typical ensemble techniques. Among which, Bagging and Boosting uses the same classification algorithm for each individual classifier. The difference between two methods is bagging trains each individual classifier independently while boosting trains dependently. Stacking constructs an ensemble on different classification algorithms, and combines multiply results with voting.

We propose a selective ensemble learning method. At first, training multiply different individual classifiers $C = \{c_1, c_2, \dots, c_n\}$ on each LCP , and evaluating the classification accuracy $CA = \{CA_1, CA_2, \dots, CA_n\}$ of each classifier c_i . Secondly, according CA_i , the method selects the classifiers $C' = \{c_1, c_2, \dots, c_t\} (t < n)$ with better performance on each LCP . Thirdly, considering the incomplete rate, the method combines the suitable classifiers of all local complete data based on the voting matrix (VM) to replace the voting policy.

Definition 4. *Incomplete rate* (α) is defined as $\alpha = \frac{|CFS| \times |U'|}{|F| \times |U|}$, where $|\bullet|$ denotes the number of objects in set, and then α_j stands for the incomplete rate of j th local complete pattern.

Definition 5. According to the definition of α_j , *voting matrix* $VM = [\alpha_1 \times VM_1 \ \alpha_2 \times VM_2 \ \dots \ \alpha_k \times VM_k]$ ($j = 1, 2, \dots, k$) determines the class label of a new case. $VM_j = \max_{i=1}^n (\sum_{k=1}^m g_{ki})$ represents the classification result of j th LCP , where m means the number of classifier, and n means the number of class label. The entry g_{ki} stands for the classification accuracy of k th classifier forecast that the new case belongs to the i th class label. If the k th classifier predicts the new case is the i th class label, g_{ki} is equal to the corresponding classification accuracy and zero otherwise.

2.3 Classify a New Case

When classifying a new case x , the relationship between the complete feature set CFS' of x and CFS of the training data set should be considered. If $CFS' \subseteq CFS$, the voting matrix can be used to classify the new case directly. If $CFS' \not\subseteq CFS$, we utilize the mutual information between the feature and class label to find the nearest neighbor complete mode, and then classify the new case.

Considering **Example 1**, a new instance has the feature value $f1=1, f2=*, f3=1, f4=*, f5=1$. The complete feature set of the new instance is $CFS' = \{f1, f3, f5\}$. From **Example 1**, we can deduce $CFS' \not\subseteq CFS$. Based on mutual information, the nearest neighbor of CFS' is $CFS3$. We classify the new instance on the best classifier of $CFS3$, and deducing the class label of the instance is $y = 1$.

3 Experimental Results

To evaluate the performance of the proposed method, we perform the experiments on three public datasets: IRIS, Lymphography, and SPECT Heart, which can be downloaded at <http://archive.ics.uci.edu/ml/>. Because these three datasets are complete, we can evaluate the classification performance effectively.

In each dataset, we replace randomly the feature value of an instance with * to construct the incomplete dataset with different α defined as **Definition 4**. In all the experiments, we set the parameter $\alpha = \{0, 0.95, 0.90, 0.80\}$, and select the individual classifiers $C = \{BayesNet, NaiveBayes, C4.5, RBFNetwork, SVM\}$. In order to compare the performance of these three methods: deletion, imputation, and **SECID**, for each of them, we perform 10 replicates of a 10-fold cross validation experiment on each data set. The average classification accuracies on each data set are listed in Table 3. The best result of classification on each data set is emphasized using a boldfaced font.

Table 3. The performance comparison in three public dataset

<i>IR</i>	Classification accuracy (%)			<i>IR</i>	Classification accuracy (%)		
	Deletion	Imputation	SECID		Deletion	Imputation	SECID
0	96	96	96	0	85.14	85.14	85.14
0.95	95.77	93.33	95.33	0.95	85.82	83.78	84.46
0.90	91.30	92.67	94.67	0.90	81.21	81.08	84.46
0.80	88.67	85.67	90.33	0.80	78.92	77.43	80.41

<i>IR</i>	Classification accuracy (%)		
	Deletion	Imputation	SECID
0	82.40	82.40	82.40
0.95	82.28	82.02	82.02
0.90	82	80.52	82.40
0.80	80.22	78.08	80.90

(a) Iris Dataset

(b) Lymphography Dataset

(c) SPECT Heart Dataset

From Table 3 it can be seen that the performance of **SECID** is lower than that of deletion with $\alpha = 0.95$, while performs well on three datasets with $\alpha = 0.9$ and $\alpha = 0.8$. This shows that the performance of **SECID** is more stable than other two methods, and the effect will be more obvious with the number increment of the incomplete instances.

4 Conclusions

Although various classification algorithms have been proposed, most of them deal with complete data. In fact, incomplete data are avoidable in real world. In this paper, we present a selective ensemble approach **SECID** for classifying incomplete data. By

mining the complete data subsets from the incomplete data, the proposed approach solves the problem of constructing classification model under incomplete decision system. Our method avoids the shortage of imputation and deletion method, and needs no assumption about the *MAR* incomplete mechanism. Experiments on three datasets show that the performance of *SECID* is more stable than that of deletion and imputation.

Acknowledgments. This research is supported by the National Science and Technology Major Special Project of China on Water Environmental Risk Assessment and Early Warning Demonstration Platform about Three Gorges (Grant No. 2009ZX07528-003).

References

1. Chen, J., Huang, H., Tian, F., Tian, S.: A Selective Bayes Classifier for Classifying Incomplete Data Based on Gain Ratio. *Knowledge-Based Systems* 21, 530–534 (2008)
2. Quinlan, J.R.: *C4.5 Programs for Machine Learning*. Morgan Kaufman, San Francisco (1993)
3. Qu, H., Mao, L., Wang, J.: Method for Self-Extracting Diagnostic Rules of Blood Stasis Syndrome Based on Decision Tree. *Chinese Journal of Biomedical Engineering* 24, 709–711, 727 (2005)
4. Boujelben, M.A., Smet, Y.D., Frikha, A., Chabchoub, H.: Building A Binary Outranking Relation in Uncertain, Imprecise and Multi-experts Contexts: The application of Evidence Theory. *International Journal of Approximate Reasoning* 50, 1259–1278 (2009)
5. Salama, A.S.: Topological Solution of Missing Attribute Values Problem in Incomplete Information Tables. *Information Sciences* 180, 631–639 (2010)
6. Huang, Z., Li, J., Su, H., Watts, G.S., Chen, H.: Large-scale Regulatory Network Analysis from Microarray Data: Modified Bayesian Network Learning and Association Rule Mining. *Decision Support Systems* 43, 1207–1225 (2007)
7. Williams, D., Liao, X., Xue, Y., Carin, L., Krishnapuram, B.: On Classification with Incomplete Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 427–436 (2007)
8. Ramoni, M., Sebastiani, P.: Robust Bayes Classifiers. *Artificial Intelligence* 125, 209–226 (2001)
9. Partalas, I., Tsoumakas, P.G., Vlahavas, I.: Pruning an Ensemble of Classifiers via Reinforcement Learning. *Neurocomputing* 72, 1900–1909 (2009)
10. Breiman, L.: Bagging Predictors. *Machine Learning* 24, 123–140 (1996)
11. Freund, Y., Schapire, R.E.: A Decision-Theoretic Generalize of On-Line Learning and An Application to Boosting. *Journal of Computer and System Sciences* 55, 119–139 (1997)
12. Witten, I.H., Frank, E.: *Data Mining*. Morgan Kaufmann Publishers, Elsevier (2005)

Hesitant Fuzzy Linguistic Term Sets

Rosa M. Rodríguez¹, Luis Martínez¹, and Francisco Herrera²

¹ Computer Science Department, University of Jaén, 23071 - Jaén, Spain
rrodrig@ujaen.es, martin@ujaen.es

² Computer Science and A.I. Department, University of Granada, 18071 - Granada, Spain
herrera@decsai.ugr.es

Abstract. Dealing with vague or imprecise information has been always a challenging problem. Different tools have been proposed to manage that uncertainty. A new model based on hesitant fuzzy sets was presented to manage situations where experts hesitate among several values to assess alternatives, variables, etc. Hesitant fuzzy sets models quantitative settings, however, it could occur similar situations but in qualitative settings, where experts think of several possible linguistic values or richer expressions than a single linguistic term to assess alternatives, variables, etc. In this contribution the aim is to introduce the concept of *Hesitant Fuzzy Linguistic Term Sets* (HFLTS) that will provide a linguistic elicitation based on the fuzzy linguistic approach and the use of context-free grammars.

Keywords: Hesitant fuzzy sets, fuzzy linguistic approach, context-free grammar, linguistic information.

1 Introduction

Problems defined in context with uncertainty are quite common in real world, but very challenging due to the difficulty to model and deal with such an uncertainty. Different tools have been proposed to solve those problems, however, there are situations where the uncertainty is not probabilistic in nature, but rather imprecise or vague. Other models as fuzzy logic and fuzzy sets theory [28] have been successfully applied to manage vague and imprecise information [18]. The modeling tools of ordinary fuzzy sets are limited when two or more sources of vagueness appear simultaneously. For this reason, different generalizations and extensions of fuzzy sets have been introduced:

- *Type 2 fuzzy sets* [5,17], and *type n fuzzy sets* [5] that include uncertainty about the membership function in their definition.
- *Intuitionistic fuzzy sets* [1] that extends fuzzy sets by an additional uncertainty degree.
- *Fuzzy multisets* [26] based on multisets that allow elements repeated in the set.
- *Hesitant fuzzy sets* proposed by Torra [23] that try to manage those situations where a set of values are possible in the definition process of the membership of an element.

The previous fuzzy tools suit problems defined in quantitative context, but sometimes, the uncertainty is due to the vagueness of meaning used by experts in problems whose

nature is rather qualitative. In such situations, the fuzzy linguistic approach [29] has provided good results in different fields and applications [11][12][13][19]. However, the use of fuzzy linguistic approach also presents some limitations mainly regarding information modeling and computational processes, called processes of *computing with words* (CW) [10][14][16]. Different linguistic models have been proposed to extend and improve the fuzzy linguistic approach:

- *The linguistic model based on type-2 fuzzy sets* representation [15][24] that represents the semantics of the linguistic terms by type-2 membership functions and use interval type-2 fuzzy sets for CW.
- *The linguistic 2-tuple model* [6] that keeps the accuracy in the processes of CW by means of a parameter, so-called *symbolic translation*.
- *The proportional 2-tuple model* [25] that generalizes and extends the 2-tuple model by using two linguistic terms with their proportion to model more accurately the information and perform the processes of CW.
- Other extensions based on the previous ones were presented in [4][9].

Revising the fuzzy linguistic approach, different linguistic extensions and generalizations, it is observed that the modeling of linguistic information is still limited because experts provide their assessments by using single and simple terms over alternatives, variables, etc. However, it might occur that experts are thinking of several linguistic terms at the same time or looking for a more complex linguistic term that are not defined in the linguistic term set.

Therefore, to overcome such limitations and taking into account the concept of hesitant fuzzy sets provided by Torra [23] to deal with several values in a membership function in a quantitative setting, in this paper we present the concept of HFLTS based on the fuzzy linguistic approach that will serve as basis to increase the flexibility of the elicitation of linguistic information. Additionally, different operations and properties of HFLTS are introduced. Afterwards, it is presented their use to improve the elicitation of linguistic information by using the fuzzy linguistic approach and context-free grammars.

The paper is organised as follows: Section 2 reviews some basic concepts necessary to understand easily our proposal. Section 3 defines the concept of HFLTS and different properties and operations to carry out processes of CW. Section 4 presents the use of HFLTS to facilitate and increase the flexibility to elicit the linguistic information. Finally, Section 5 points out some concluding remarks and future works.

2 Preliminaries

In this section, we review briefly some concepts of the fuzzy linguistic approach [29] and hesitant fuzzy sets [23] to understand the proposal of HFLTS and its use.

2.1 Fuzzy Linguistic Approach

In many real decision situations are suitable the use of linguistic information due to the nature of the problem. In these cases, the fuzzy linguistic approach [29] models the

linguistic information by using the fuzzy set theory [28] to manage the uncertainty and model the information.

Zadeh [29] introduced the concept of linguistic variable as *a variable whose values are not numbers but words or sentences in a natural or artificial language*. A linguistic value is less precise than a number, but it is closer to the natural language used by human beings.

To deal with linguistic variables, it is necessary to choose the appropriate linguistic descriptors for the linguistic term sets and their semantics. To do so, there are different possibilities [27][29]. The choice of the linguistic descriptors can be carried out as follows:

- Supplying directly the term set by considering all the terms symmetrically distributed on a scale which has an order defined [27]. In these cases, it is usually necessary that exist the following operators: (i) Negation: $\text{Neg}(s_i) = s_j$ such that $j = g - i$ ($g + 1$ is the cardinality), (ii) Maximization: $\max(s_i, s_j) = s_i$ if $s_i \geq s_j$, (iii) Minimization: $\min(s_i, s_j) = s_i$ if $s_i \leq s_j$.
- Defining the linguistic term set by means of a context-free grammar, G , such that the linguistic terms are sentences generated by G [2][3][29]. A grammar G is a 4-tuple (V_N, V_T, I, P) , where V_N is the set of non-terminal symbols, V_T is the set of terminals symbols, I is the starting symbol, and P the production rules defined in an extended Backus Naur Form [3]. Among the terminal symbols of G , we can find primary terms (e.g., low, medium), hedges (e.g., not, very), relations (e.g., lower than, higher than), conjunctions (e.g., and, but), and disjunctions (e.g., or).

And the definition of their semantics can be accomplished as [27][29]:

- A semantics based on membership functions and a semantic rule. It assumes that the meaning of each linguistic term is given by means of a fuzzy subset defined in the interval $[0,1]$, which is described by membership functions [3]. This semantic approach is used when the linguistic descriptors are generated by means of a context-free grammar.
- A semantics based on an ordered structure of the linguistic term set that introduces the semantics from the structure defined over the linguistic term set. So, the users use an ordered linguistic term set to provide their assessments [22][27].
- Mixed semantics that uses elements from the previous approaches.

2.2 Hesitant Fuzzy Sets

Torra presented in [23] the definition of hesitant fuzzy sets to fulfil the management of decision situations in quantitative contexts where the decision makers hesitate among different possible values to assess an alternative or criterion.

A hesitant fuzzy set is defined in terms of a function that returns a set of membership values for each element in the domain [23]:

Definition 1. [23] *Let X be a reference set, a hesitant fuzzy set on X is a function h that returns a subset of values in $[0,1]$.*

$$h : X \rightarrow \{[0, 1]\}$$

A hesitant fuzzy set can be also defined in terms of the union of their membership degree to a set of fuzzy sets.

Definition 2. [23] Let $M = \{\mu_1, \mu_2, \dots, \mu_n\}$ be a set of n membership functions. The hesitant fuzzy set associated with M , h_M , is defined as:

$$h_M : M \rightarrow \{[0, 1]\}$$

$$h_M(x) = \bigcup_{\mu \in M} \{\mu(x)\}$$

Some basic operations with hesitant fuzzy sets, such as, union, intersection, complement and so on were defined in [23].

3 Hesitant Fuzzy Linguistic Term Sets

In qualitative contexts might occur that experts hesitate among several linguistic values. As it was pointed out in the introduction several proposals have been proposed in the literature [6,25]. However, all of them are still limited and are not adequate to fulfil the necessities and requirements of experts in hesitant situations.

In this section is introduced the concept of HFLTS based on the fuzzy linguistic approach and hesitant fuzzy sets. Additionally, some basic operations of HFLTS are defined.

3.1 Concept and Basic Operations

Definition 3. Let S be a linguistic term set, $S = \{s_0, \dots, s_g\}$, a HFLTS, H_S , is an ordered finite subset of consecutive linguistic terms of S .

Let S be a linguistic term set, $S = \{s_0, \dots, s_g\}$, we then define the empty and full HFLTS for a linguistic variable, x , as follows:

- Empty HFLTS: $H_S(x) = \{\}$
- Full HFLTS: $H_S(x) = S$

Any other HFLTS is formed at least with one linguistic term in S .

Example 1. Let S be a linguistic term set, $S = \{s_0 : \text{nothing}, s_1 : \text{very_low}, s_2 : \text{low}, s_3 : \text{medium}, s_4 : \text{high}, s_5 : \text{very_high}, s_6 : \text{perfect}\}$, different HFLTS might be:

$$H_S(x) = \{\text{very_low}, \text{low}, \text{medium}\}$$

$$H_S(x) = \{\text{high}, \text{very_high}, \text{perfect}\}$$

Once defined the concept of HFLTS, it is necessary to introduce operations and computations that can be performed on them.

Let S be a linguistic term set, $S = \{s_0, \dots, s_g\}$ and H_S, H_S^1 , and H_S^2 three HFLTS:

Definition 4. The upper bound, H_{S+} , and lower bound, H_{S-} , of the HFLTS, H_S , are defined as:

- $H_{S+} = \max(s_i, s_j) = s_i$, if $s_i \geq s_j$; $s_i, s_j \in H_S$
- $H_{S-} = \min(s_i, s_j) = s_i$, if $s_i \leq s_j$; $s_i, s_j \in H_S$

Definition 5. The complement of HFLTS, H_S , is defined as:

$$H_S^c = S - H_S = \{s_i / s_i \in S \text{ and } s_i \notin H_S\}$$

Proposition 1. The complement of a HFLTS is involutive:

$$(H_S^c)^c = H_S$$

Definition 6. The union between two HFLTS, H_S^1 and H_S^2 is defined as:

$$H_S^1 \cup H_S^2 = \{s_i / s_i \in H_S^1 \text{ or } s_i \in H_S^2\}$$

the result will be another HFLTS.

Definition 7. The intersection of two HFLTS, H_S^1 and H_S^2 is:

$$H_S^1 \cap H_S^2 = \{s_i / s_i \in H_S^1 \text{ and } s_i \in H_S^2\}$$

the result of this operation is another HFLTS.

The comparison between linguistic terms is necessary in many problems and has been defined in different approaches. The comparison between HFLTS is not simple, therefore, we introduce the concept of envelope of a HFLTS.

Definition 8. The envelope of the HFLTS, $env(H_S)$, is a linguistic interval whose limits are obtained by means of upper bound (max) and lower bound (min), hence:

$$env(H_S) = [H_{S-}, H_{S+}], \quad H_{S-} \leq H_{S+}$$

Example 2. Let $S = \{\text{nothing, very_low, low, medium, high, very_high, perfect}\}$ be a linguistic term set, and $H_S = \{\text{very_low, low, medium}\}$ be a HFLTS of S , its envelope is:

$$H_{S-}(\text{very_low, low, medium}) = \mathbf{very_low}, \quad H_{S+}(\text{very_low, low, medium}) = \mathbf{medium}$$

$$env(H_S) = [\text{very_low, medium}]$$

Definition 9. The definition of the comparison between two HFLTS is based on the concept of envelope of the HFLTS, $env(H_S)$. Hence, the comparison between, H_S^1 and H_S^2 is defined as follows:

$$H_S^1(x) > H_S^2(x) \text{ if } env(H_S^1(x)) > env(H_S^2(x))$$

$$H_S^1(x) = H_S^2(x) \text{ if } env(H_S^1(x)) = env(H_S^2(x))$$

As consequence the comparison is carried out by interval values. In the literature have been introduced different approaches to comparing intervals [7][8][21]. We use the approach presented by Sengupta in [21] to accomplish the comparison of HFLTS, more detail can be found in [20].

3.2 Properties

To conclude this section some important properties of the HFLTS operations are reviewed.

Let H_S^1 , H_S^2 and H_S^3 be three HFLTS and $S = \{s_0, \dots, s_g\}$, then:

– *Commutativity*

$$H_S^1 \cup H_S^2 = H_S^2 \cup H_S^1$$

$$H_S^1 \cap H_S^2 = H_S^2 \cap H_S^1$$

– *Associative*

$$H_S^1 \cup (H_S^2 \cup H_S^3) = (H_S^1 \cup H_S^2) \cup H_S^3$$

$$H_S^1 \cap (H_S^2 \cap H_S^3) = (H_S^1 \cap H_S^2) \cap H_S^3$$

– *Distributive*

$$H_S^1 \cap (H_S^2 \cup H_S^3) = (H_S^1 \cap H_S^2) \cup (H_S^1 \cap H_S^3)$$

$$H_S^1 \cup (H_S^2 \cap H_S^3) = (H_S^1 \cup H_S^2) \cap (H_S^1 \cup H_S^3)$$

Due to the long limitation of the paper, the demonstrations of these properties can be found in [20].

4 Elicitation of Information Linguistic Based on HFLTS

The main objective of the definition of HFLTS is to improve the flexibility of the elicitation of linguistic expressions when experts hesitate among several linguistic values to assess alternatives or criteria.

So far, it has been introduced the concept of HFLTS that can be directly used by the experts to elicit several linguistic values for a linguistic variable, but such elements are not similar to the human beings way of thinking and reasoning. Therefore, in this section it is proposed the definition of linguistic sentences that are more similar to human beings expressions and semantically represented by means of HFLTS and generated by a context-free grammar.

Definition 10. Let G_H be a context-free grammar and $S = \{s_0, \dots, s_g\}$ a linguistic term set. The elements of $G_H = (V_N, V_T, I, P)$ are defined as follows:

$$V_N = \{ \langle \text{primary term} \rangle, \langle \text{composite term} \rangle, \langle \text{unary relation} \rangle, \langle \text{binary relation} \rangle, \langle \text{conjunction} \rangle \}$$

$$V_T = \{ \text{lower than}, \text{greater than}, \text{between}, \text{and}, s_0, s_1, \dots, s_g \}$$

$$I \in V_N$$

The production rules are defined in an extended Backus Naur Form such that the brackets enclose optional elements and the symbol | indicate alternative elements [3]. For the context-free grammar, G_H , the production rules are the following ones:

$P = \{I ::= \langle \text{primary term} \rangle | \langle \text{composite term} \rangle$
 $\langle \text{composite term} \rangle ::= \langle \text{unary relation} \rangle \langle \text{primary term} \rangle | \langle \text{binary relation} \rangle$
 $\langle \text{primary term} \rangle \langle \text{conjunction} \rangle \langle \text{primary term} \rangle$
 $\langle \text{primary term} \rangle ::= s_0 | s_1 | \dots | s_g$
 $\langle \text{unary relation} \rangle ::= \text{lower than} | \text{greater than}$
 $\langle \text{binary relation} \rangle ::= \text{between}$
 $\langle \text{conjunction} \rangle ::= \text{and} \}$

Remark 1. *The unary relation has some limitations. If the non-terminal symbol is lower than, the primary term cannot be s_0 and if the non-terminal symbol is greater than the primary term cannot be s_g .*

Remark 2. *In the binary relation the primary term of the left side must be less than the primary term of the right side.*

Example 3. *Let $S = \{ \text{nothing, very_low, low, medium, high, very_high, perfect} \}$ be a linguistic term set, some linguistic expressions obtained by means of the context-free grammar, G_H , might be:*

- $ll_1 = \text{very_low}$
- $ll_2 = \text{lower than low}$
- $ll_3 = \text{greater than high}$
- $ll_4 = \text{between high and very_high}$

It was also defined a transformation function, E_{G_H} , to obtain HFLTS from the linguistic expressions, ll , generated by the context-free grammar, G_H .

Definition 11. *Let E_{G_H} be a function that transforms linguistic expressions, ll , obtained by G_H , into HFLTS, H_S , where S is the linguistic term set used by G_H .*

$$E_{G_H} : ll \longrightarrow H_S$$

The linguistic expressions generated by using the production rules are transformed into HFLTS in different ways as follows:

- $E_{G_H}(s_i) = \{s_i / s_i \in S\}$
- $E_{G_H}(\text{less than } s_i) = \{s_j / s_j \in S \text{ and } s_j \leq s_i\}$
- $E_{G_H}(\text{greater than } s_i) = \{s_j / s_j \in S \text{ and } s_j \geq s_i\}$
- $E_{G_H}(\text{between } s_i \text{ and } s_j) = \{s_k / s_k \in S \text{ and } s_k \geq s_i \text{ and } s_k \leq s_j\}$

5 Conclusions and Future Works

This contribution has introduced the concept of HFLTS to increase the flexibility and richness of linguistic elicitation based on the fuzzy linguistic approach and the use of context-free grammars to support the elicitation of linguistic information by experts in hesitant situations in qualitative contexts. Additionally, different operations and properties of HFLTS have been presented.

In the future, it will be studied the application of HFLTS to decision making processes defined under uncertainty where experts will be able to provide their assessments by using linguistic expressions based on HFLTS similar to the expressions used by human beings.

Acknowledgements. This work is partially supported by the Research Project TIN-2009-08286, P08-TIC-3548 and FEDER funds.

References

1. Atanassov, K.T.: Intuitionistic fuzzy sets. *Fuzzy Sets and Systems* 20, 87–96 (1986)
2. Bonissone, P.P.: A fuzzy sets based linguistic approach: theory and applications. In: Gupta, M.M., Sanchez, E. (eds.) *Approximate Reasoning in Decision Analysis*, pp. 99–111. North-Holland Publishing Company (1982)
3. Bordogna, G., Pasi, G.: A fuzzy linguistic approach generalizing boolean information retrieval: A model and its evaluation. *Journal of the American Society for Information Science* 44, 70–82 (1993)
4. Dong, Y., Xu, Y., Yu, S.: Computing the numerical scale of the linguistic term set for the 2-tuple fuzzy linguistic representation model. *IEEE Transactions on Fuzzy Systems* 17(6), 1366–1378 (2009)
5. Dubois, D., Prade, H.: *Fuzzy Sets and Systems: Theory and Applications*. Kluwer Academic, New York (1980)
6. Herrera, F., Martínez, L.: A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems* 8(6), 746–752 (2000)
7. Ishibuchi, H., Tanaka, H.: Theory and methodology: Multiobjective programming in optimization of the interval objective function. *European Journal of Operational Research* 48, 219–225 (1990)
8. Kundu, S.: Min-transitivity of fuzzy leftness relationship and its application to decision making. *Fuzzy Sets and Systems* 86, 357–367 (1997)
9. Li, D.F.: Multiattribute group decision making method using extended linguistic variables. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 17(6), 793–806 (2009)
10. Liu, J., Martínez, L., Wang, H., Rodríguez, R.M., Novozhilov, V.: Computing with words in risk assessment. *International Journal of Computational Intelligence Systems* 3(4), 396–419 (2010)
11. Martínez, L.: Sensory evaluation based on linguistic decision analysis. *International Journal of Approximate Reasoning* 44(2), 148–164 (2007)
12. Martínez, L., Liu, J., Yang, J.B.: A fuzzy model for design evaluation based on multiple criteria analysis in engineering systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 14(3), 317–336 (2006)
13. Martínez, L., Pérez, L.G., Barranco, M.: A multi-granular linguistic based-content recommendation model. *International Journal of Intelligent Systems* 22(5), 419–434 (2007)
14. Martínez, L., Ruan, D., Herrera, F.: Computing with words in decision support systems: An overview on models and applications. *International Journal of Computational Intelligence Systems* 3(4), 382–395 (2010)
15. Mendel, J.M.: An architecture for making judgement using computing with words. *International Journal of Applied Mathematics and Computer Sciences* 12(3), 325–335 (2002)
16. Mendel, J.M., Zadeh, L.A., Yager, R.R., Lawry, J., Hagsras, H., Guadarrama, S.: What computing with words means to me. *IEEE Computational Intelligence Magazine* 5(1), 20–26 (2010)
17. Mizumoto, M., Tanaka, K.: Some properties of fuzzy sets of type 2. *Information Control* 31, 312–340 (1976)
18. Parsons, S.: Current approaches to handling imperfect information in data and knowledge bases. *IEEE Transactions on Knowledge Data Engineering* 8(3), 353–372 (1996)

19. Rodríguez, R.M., Espinilla, M., Sanchez, P.J., Martínez, L.: Using linguistic incomplete preference relations to cold start recommendations. *Internet Research* 20(3), 296–315 (2010)
20. Rodríguez, R.M., Martínez, L., Herrera, F.: Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems* (2011), doi:10.1109/TFUZZ.2170076
21. Sengupta, A., Kumar Pal, T.: On comparing interval numbers. *European Journal of Operational Research* 127, 28–43 (2000)
22. Torra, V.: Negation functions based semantics for ordered linguistic labels. *International Journal of Intelligent Systems* 11, 975–988 (1996)
23. Torra, V.: Hesitant fuzzy sets. *International Journal of Intelligent Systems* 25(6), 529–539 (2010)
24. Türkşen, I.B.: Type 2 representation and reasoning for CWW. *Fuzzy Sets and Systems* 127, 17–36 (2002)
25. Wang, J.H., Hao, J.: A new version of 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems* 14(3), 435–445 (2006)
26. Yager, R.R.: On the theory of bags. *International Journal Generation System* 13, 23–37 (1986)
27. Yager, R.R.: An approach to ordinal decision making. *International Journal of Approximate Reasoning* 12(3-4), 237–261 (1995)
28. Zadeh, L.: Fuzzy sets. *Information and Control* 8, 338–353 (1965)
29. Zadeh, L.: The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences, Part I, II, III* (8,9), 199–249, 301–357, 43–80 (1975)

A Conceptual Model for Risk-Based Situation Awareness

Mohsen Naderpour¹, Jie Lu¹, and Etienne Kerre²

¹ Decision Systems and e-Service Intelligence Laboratory
Centre for Quantum Computation and Intelligent Systems (QCIS)
School of Software, Faculty of Engineering and Information Technology
University of Technology, Sydney, P.O. Box 123, Broadway, NSW 2007, Australia
Mohsen.Naderpour@student.uts.edu.au, Jie.Lu@uts.edu.au

² Department of Applied Mathematics & Computer Science
Ghent University, B-9000 Gent, Belgium
Etienne.Kerre@ugent.be

Abstract. Situation Awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future. It is a crucial factor in decision-making in a dynamic environment particularly with certain degrees of risk, called risk-based situation awareness. In this paper we first explore the most popular models in situation awareness, data fusion and risk assessment. We show how they complement each other in developing a conceptual model for risk-based situation awareness. We will also demonstrate how this model can be used to support decision-making in a dynamic environment.

Keywords: Situation Awareness, Fusion Model, Decision Making, Dynamic Risk Assessment.

1 Introduction

Literature and real world applications have reported that the human error is the biggest challenge within most industries and the most likely cause of an accident. It implies that people are merely careless or poorly trained or somehow not very reliable in general. In fact, in the vast majority of these accidents the human operator was striving against significant challenges. They have to face both data overload and the challenge of working with a complex system. They are drilled with long lists of procedures and checklists designed to cope with some of these difficulties, but from time to time they are apt to fail. In [1] the authors have presented an intelligent reduction method to reduce the huge number of control variables to the most relevant ones in case of an emergency. In fact, the person is not the cause of these errors but is the final dumping ground for the inherent problems and difficulties in the technologies we have created. Operators generally have no difficulty in performing their tasks physically, and no difficulty in knowing what is the correct thing to do, but they are stressed by the task of understanding what is going on in the situation. So in the last two decades, lots of researches have been done about Situation Awareness (SA) and many frameworks have been proposed for building SA. However, the most commonly used models are Endsley's model [5] and the Joint Directors of Laboratories (JDL) fusion model [2].

Endsley's view is based on cognitive principles. It is primarily a mental model and divided into three major components. On the other hand, the JDL fusion model provides a functional data centric approach that is particularly suited for the tactical domain and divides into five levels [4]. In complex systems, level 1 situation awareness is highly supported through the various heterogeneous sensors and appropriate signal-processing methods for extracting as much information as possible about the dynamic environment and its elements, but regarding levels 2 and 3 situation awareness, there is still a need for concepts and methods supporting them that are able to infer real situations and to project their status in the near future [12]. Especially, there is no framework that assesses situations and helps the decision maker to take appropriate action in hazardous situations. Therefore, we propose an innovative conceptual model which aims to focus on this.

The next section gives a brief review of basic concepts and related work. In the Section 3 we present a conceptual model and describe its components. In Section 4 we describe a functional demonstration of the proposed model. The paper finishes with a conclusion and future work in Section 5.

2 Basic Concepts and Related Work

2.1 Situation Awareness

During the operation of complex systems that include human decision making, acquiring and interpreting information from the environment forms the basis for the state of knowledge of a decision maker. This state is often referred to as situation awareness. One of the earliest and most widely applicable Situation Awareness (SA) definitions describes it as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” [5]. Therefore SA involves perceiving critical factors in the environment (Level 1), understanding what those factors signify (Level 2), and anticipating what will happen with the situation in the near future (Level 3; see Figure 1). Endsley notes that SA is particularly relevant in work domains where information needs to be processed very quickly and the consequences of poor decisions can be serious [11].



Fig. 1. Three levels of situation awareness

2.2 Joint Directors of Laboratories (JDL) Fusion Model

A lot of data fusion models have been developed and compared to Endsley's situation awareness model, whereas the most common model is the JDL fusion process model. As shown in Figure 2, the JDL model has five levels. A stream of data enters the model at level 0, Sub-object Data Assessment. This level provides physical access to the raw bits or signal. In addition, estimation and prediction of the existence of an object is performed based on pixel or signal level data association and characterization. Objects are correlated and tagged over time in an attempt to build tracks and to perform object identification during level 1 processing, or Object Assessment. During Situation Assessment, or level 2 processing, the knowledge of objects, their characteristics, relationships with each other and cross force relations are aggregated in an attempt to understand the current situation. After Situation Assessment, the impact of the given situation must be assessed (Level 3-Impact Assessment). The final level, Process Refinement, provides a feedback mechanism to each of the other layers, including the sensor itself [4].

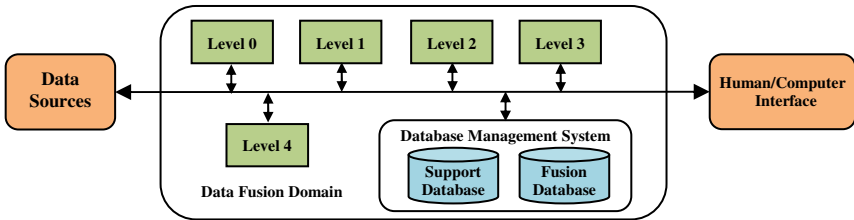


Fig. 2. JDL fusion model

2.3 Risk Assessment

There are lots of conventional methods and techniques for risk assessment which are proposed as a backbone to the identification, specification, evaluation and assessment of the undesirable events or properties adversely affecting technical functionality, cost, reliability, safety, quality etc. The common risk assessment process comprises seven systemic aspects. Hazards are all potential objects, conditions or states which are likely to lead to loss at any stage/phase or mode of the system. When hazards are identified and recorded, it is often desirable to understand their causality. This may help in eliminating the hazardous condition altogether to take actions to reduce its likelihood or frequency. It is also desirable to develop an understanding picture of incidents and accidents that are likely to arise from a hazard. This is the third principle, consequence analysis. It is also necessary to establish the degree of loss that each forecast accident may cause. So loss will be an aggregation of each accident's safety, commercial and environmental effects. The estimation of likelihood/frequency and the associated losses pave the way to understanding the risks associated with a given hazard. When the risks are evaluated, it would be desirable to establish a set of practical options to reduce them and estimate their effectiveness and establish a decision criterion based on the costs and the expected benefits from each option [9].

2.4 Related Work

Situation awareness involves a variety of cognitive processes employed by operators in a complex and dynamic environment to understand the current state of the environment in order to anticipate its future state. This concept was first recognized in solving problems for crews in military aircraft during the World War I [5]. In the mid-1970s the US military ergonomists started investigating the factors affecting aircrew, and from then onwards, SA became an established concept. The concept was later adopted by human factors researchers for studies of complex environments such as air traffic controllers, nuclear power plant operators, anesthesiologists, military commanders, electronic warfare tacticians, automobile drivers, and so on [5]. McGuinness and Foy extended Endsley's Model by adding a fourth level, which is called Resolution. This level provides awareness of the best path to follow to achieve the desired outcome to the situation [4]. There are numerous examples of accidents caused by a failure of SA, for example driving accidents [3] and operational errors in air traffic control [5]. Roy proposed a situation analysis process to provide and maintain a state of situation awareness. Salerno and his colleagues have proposed a framework for situation awareness under the title of data fusion [2] and recently a novel approach by emphasizing the role of formal logics and automated reasoning system has been built [10]. The concept of situation management in dynamic systems proposed by Jakobson includes not only the processes of perceiving and recognizing situations, but also the analysis of past situations and the prediction of future situations [7]. A rough taxonomy of functions related to situation assessment is proposed by Steinberg and a general overview of current approaches to automating this process is given [6].

3 A Risk-Based Situation Awareness Conceptual Model

The concept of SA established by Endsley [5] is applicable in many different domains and it can also be used for dynamic risk assessment. We propose a conceptual model in order to help the decision makers to know and assess the situation according to risk assessment principles and adopt the appropriate actions for eliminating or reducing the unacceptable and intolerable risks.

For determining the aspects of the situation that are important for an operator's SA, we use the goal-directed task analysis. The results are showed in Table 1. In this analysis, the major goal and the necessary subgoals are identified. Associated with each subgoal, the major decisions that need to be made are then identified.

3.1 Perception

The first step of the conceptual model is data gathering. Data should be collected from various sensors which are distributed in intended area. The aim of this section is to locate and identify objects. For this purpose we rely on the levels 0 and 1 of the JDL model. These levels provide a global picture of the situation which is reported by fusing the attributes of an object from multiple sources. Level 2 of JDL model attempts to construct a picture from incomplete information provided by level 1, which is to relate the reconstructed entity with an observed event. Any variety of relations- physical, organizational, informational, and perceptual- can be considered, as appropriate

to the given information system's mission. Knowledge and experience of the past can be applied to the determination of the potential hazardous circumstances. Therefore it is essential to compile and maintain a database, in order to facilitate identification of the likely problem spots and where possible, avoid the errors, failures and losses of the past. This approach is seldom adequate, especially when there are novelties or significant changes in the system and there are some limitations in computers multi-media data processing. So we must use decision maker's knowledge as well. The outcome of this step is the identifying of hazardous entities or relations which have a potential to lead to loss. The results should be stored in the database to provide the elements and quantities which can be used for knowledge discovery and next analysis. As can be seen from the Figure 3, the above forms what we defined as the perception level of situation awareness.

Table 1. Safety goals and decisions

1.0	Eliminate or reduce the risks to a level that is as low as reasonably practicable
1.1	Determine the risks
1.1.1	Hazards identification <ul style="list-style-type: none"> • <i>Events or their relations have potential to damage?</i>
1.1.2	Likelihood determination <ul style="list-style-type: none"> • <i>The hazard causality?</i>
1.1.3	Severity determination <ul style="list-style-type: none"> • <i>The hazard consequences?</i> • <i>Degree of losses?</i>
1.2	Reduce the risks
1.2.1	Establish the practical options <ul style="list-style-type: none"> • <i>Reduction options (RO)?</i> • <i>Containment options (CO)?</i>
1.2.2	Impact of the options <ul style="list-style-type: none"> • <i>Impact of RO and CO options?</i>

3.2 Comprehension

The comprehension phase involves Casual, Consequence and Loss Analysis. The causal analysis techniques are predominately applied within reliability engineering and are generally supported by mathematical foundations and a suite of computer based tools. The quantification of causal models entails an objective assessment of the potential frequency or likelihood for the causal factors. These are combined according to the rules of probability calculus and Boolean logic to generate a normalized or absolute measure for the realization of the hazards. Consequence analysis is concerned about what may potentially follow the occurrence of a hazardous situation. If a hazardous situation is experienced and knowledge of its rate or likelihood of occurrence is at hand, consequence analysis would prove more beneficial than its causal counterpart in establishing the likely consequences and extent of potential losses. Loss analysis comprises the systematic investigation of the adverse outcome associated with all incidents and accidents identified through consequence analysis. Loss includes various degrees of harm to people, commercial/operational detriment to an enterprise or damage to the ecology of the environment or a combination thereof. It is useful for all three components to be converted and expressed in a common currency such as

money for potential comparison and aggregation in order to provide a coherent view of the totality of loss associated with a hazardous situation.

3.3 Projection

Depending on the consequent losses, the hazards may subsequently require risk elimination, mitigation, transfer, control or an appropriate combination thereof. So options analysis provides the future necessary actions which should be implemented to eliminate or reduce the hazards. This process corresponds to level 3 of JDL fusion model. From this step we enter to the projection phase. Upon identification and recording, it is essential to estimate the likely effects and potential benefits of each option on the consequent safety, commercial and environmental losses in order to establish the objective and systematic criteria for selection and implementation. This is a requirement of the statutory legal framework in some countries (i.e. UK) to ensure that the safety risks are reduced to As Low As Reasonably Practicable (ALARP) levels.

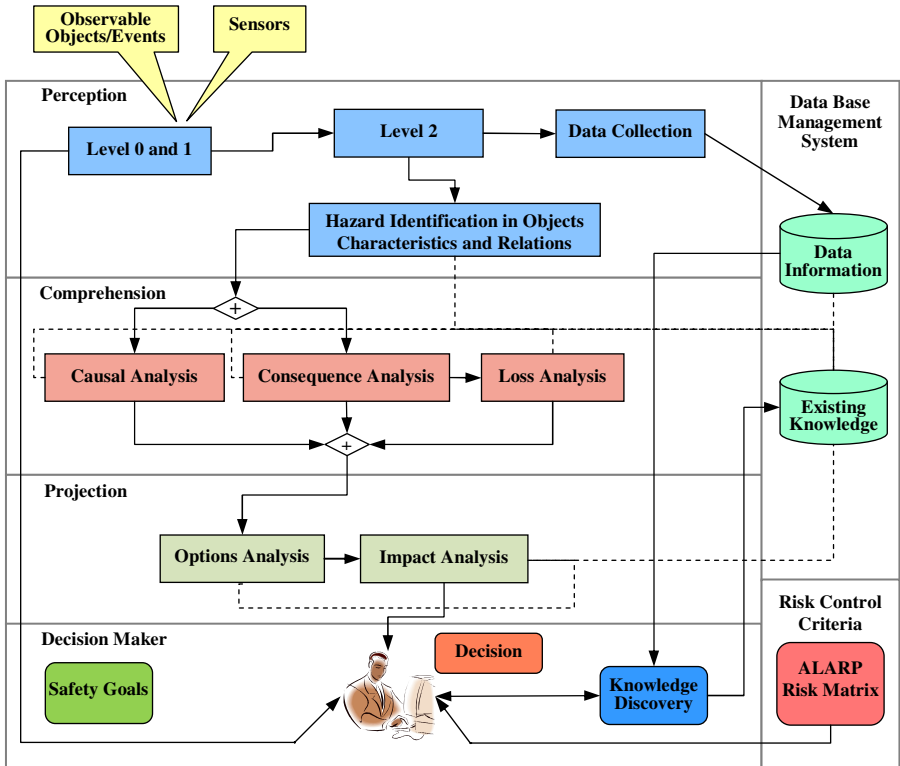


Fig. 3. The conceptual model of risk-based situation

3.4 Existing Knowledge

In order to comprehend the current objects' characteristics and their relevancy we must have some knowledge of similar situations that occurred in the past and relevant events currently occurring. If this prior knowledge does not exist, we need to learn or discover it. This knowledge can be captured as rules which can be learned by data sets and would include such concepts as casuals and consequences. This area is what we have called existing knowledge.

3.5 Knowledge Discovery

The knowledge discovery process (KDP) seeks new knowledge in some application domain. It is defined as the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data. The process consists of multiple steps that are executed sequentially. Each subsequent step is initiated upon successful completion of the previous step, and requires the result generated by the previous step as its input. Another common feature of the proposed models is the range of activities covered, which stretches from the task of understanding the project domain and data, through data preparation and analysis, to evaluation, understanding, and application of the generated results. A schematic diagram is shown in Figure 4. Typical inputs include data in various formats, such as numerical and nominal data stored in databases or flat files; images; etc. The output is the generated new knowledge, usually described in terms of rules, patterns, models, etc [8].

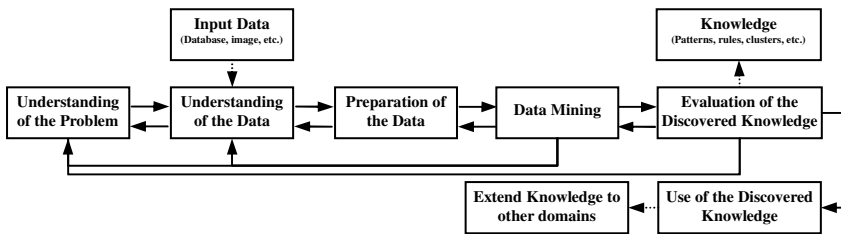


Fig. 4. Structure of knowledge discovery process

4 A Simple Functional Demonstration

In this section we attempt to describe how our conceptual model can be used in a practical environment. To achieve this goal we choose a PVC reactor unit. This process is the polymerization of vinyl chloride monomer (VCM) to make polyvinyl chloride (PVC) and includes some process instrumentations, safety facilities and equipments as shown in Figure 5 and listed in Table 2 [13].

Chemical processes are characterized by many variables. However, the experience accumulated through years by domain experts allows for the representation of behavior of chemical processes not only by the mathematical models but also by a set of rules. So a rule-based system can be used to support the operator's situational awareness. In this paper we can use three types of rules: fact, intermediate and decision

rules. The rules of the different types are stored separately in the data base. This structure ensures a quick matching between facts and fact rules.

In the perception phase we rely on the real-time data from the control system e.g. temperature and pressure and the reasoning process depends on the heuristic rules which are introduced into the knowledge base. For identifying the hazards and abnormal situations, we use fact rules. In a fact rule, the antecedent is a functional expression consisting of the principal measurement variables and the fault threshold value. An example is shown as follows:

IF TICA > 143°C; THEN high reactor temperature (hazardous situation)

In the comprehension phase we can use intermediate rules. In an intermediate rule, the antecedent is a hazard, while cause and consequent are described in its consequent. Two examples are shown as follows:

IF high reactor temperature; THEN its immediate cause is CW pump failure
IF high reactor temperature; THEN it leads to runaway reaction

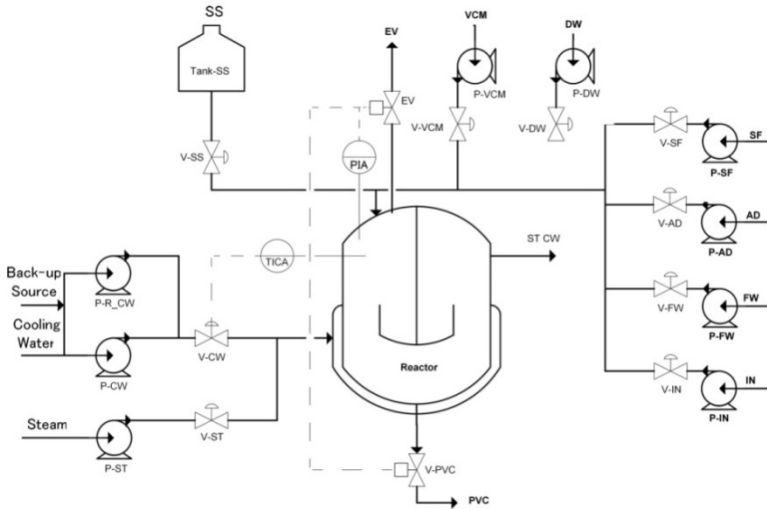


Fig. 5. PVC reactor unit

Table 2. Safety facilities for PVC process design

Designed equipments or facilities	Specific design rationale
Reactor temperature sensor	To monitor reactor temperature
Reactor pressure sensor	To monitor reactor pressure
High reactor temperature alarm	To inform high reactor temperature
High reactor pressure alarm	To inform high reactor pressure
Redundancy pump	To maintain reactor temperature in case of pump failure
Add shortstop chemical	To quench a runaway reaction
Manual valve of bottom discharge	To drop batch into a dump tank/emergency area

For determining the likelihood, according to the data base, the probability which is represent before observing the new data should be calculated. After the initiation of the process, an event which is not characterized as accident but indicates the increasing likelihood of an accident occurrence can be collected. Bayesian theory is a probabilistic approach which applies the conditional probability principles to reason with uncertainties. Considering x as the hazard probability of the system and $f(x)$ as the probability distribution function (prior distribution), $f(x|Data)$ will present the posterior distribution which is derived using the following equation [15]:

$$f(x|Data) \approx g(Data|x)f(x) \text{ where } g(Data|x) \text{ is the likelihood function}$$

The consequences of an abnormal event in chemical process industries may be categorized into 3 groups; asset loss, human fatality, environmental loss. A severity matrix can be defined and the equivalent dollar value of damage associated with each consequence category based on the severity of damage also should be determined. A sample is presented in [14]. The risk is obtained by multiplying the failure posterior probability of each hazard by its related severity. The risk level can be obtained from a risk matrix. A sample of risk matrix is described in [16]. In the projection phase we can use decision rules. A decision rule includes the corresponding suggested action. Decision rules should cover the Reduction Options and Containment Options. An example is shown as follows:

IF CW pump failure; THEN switch to redundancy pump

As a matter of fact, our rule base system consists of readable if-then statements but it cannot learn the rules itself. For discovering new knowledge, a neural network can be used to achieve readability and learning ability at the same time. The obtained rules may reveal insight into the data that generated the model, and for control purposes, they can be integrated with rules formulated by control experts (operators). Also extracting rules from data is a form of modeling activity within pattern recognition, data analysis or data mining referred to as the search for structure in data. The goal is to reduce the complexity in a problem, or to reduce the amount of data associated with a problem.

5 Conclusion and Future Work

Information fusion and situation awareness are two techniques used frequently to describe how various sources of data/information are used to build a bigger picture of the current situation. Information fusion is primarily focused on the processes of integration. Situation awareness takes these results and focuses on its representation for the end user to comprehend the current artifacts and projection of their status in the near future. We developed a risk-based situation awareness conceptual model and explained its components. What is presented here is only a starting point. One of the desired properties of SA is to establish relationships and associations among entities; it should anticipate with a priori knowledge in order to rapidly gather, assess, interpret and predict what these relationships might be; it should plan, predict, anticipate again with updated knowledge, adaptively learn, and control the fusion processes for optimum knowledge capture and decision making. Therefore our future work would be

the developing knowledge base methodologies for supporting the proposed conceptual model to have optimum SA system.

References

1. Cappelle, B., De Cooman, G., Kerre, E.E., Vanmassenhove, F.R.: Intelligent reduction methods based upon fuzzy set theory as a helpful tool for controlling industrial plants. *Fuzzy Sets and Systems* 50, 119–125 (1992)
2. Salerno, J., Hinman, M., Boulware, D.: Building a Framework for Situation Awareness. In: *The 7th International Conference on Information Fusion, Sweden* (2004)
3. Horswill, M.S., McKenna, F.P.: Drivers' hazard perception ability: Situation awareness on the road. In: Banbury, S., Tremblay, S. (eds.) *A Cognitive Approach to Situation Awareness*, Aldershot, pp. 155–172 (2004)
4. Salerno, J., Boulware, D., Cardillo, R.: Knowledge Representation Requirement for Situation Awareness. In: *8th International Conference on Information Fusion, USA* (2005)
5. Endsley, M.R.: Situation Awareness. In: Salvendy, G. (ed.) *Handbook of Human Factors and Ergonomics*, pp. 528–542. John Wiley and Sons Press (2006)
6. Steinberg, A.N.: Foundations of Situation and Threat Assessment. In: Liggins, M.E., Hall, D.L., Llinas, J. (eds.) *Handbook of Multi sensor Data Fusion: Theory and Practice*, pp. 437–501. CRC Press (2008)
7. Jakobson, G., Buford, J., Lewis, L.: Situation Management: Basic Concepts and Approaches. *Information Fusion and Geographic Information Systems*, 18–33 (2007)
8. Cios, K.J., Pedrycz, W., Swiniarski, R.W., Kurgan, L.A.: *A Knowledge Discovery Approach*, pp. 9–24. Springer press (2007)
9. Hessami, A.G.: A Systems Framework for Strategic Approach to Risk in E-Business. *International Journal of Information Science and Management, Special Issue* (2010)
10. Baader, F., Bauer, A., Baumgartner, P., Cregan, A., Gabaldon, A., Ji, K., Lee, K., Rajaratnam, D., Schwitler, R.: A Novel Architecture for Situation Awareness Systems. In: Giese, M., Waaler, A. (eds.) *TABLEAUX 2009. LNCS*, vol. 5607, pp. 77–92. Springer, Heidelberg (2009)
11. Jones, R.E.T., Connors, E.S., Endsley, M.R.: A Framework for Representing Agent and Human Situation Awareness. In: *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, USA* (2011)
12. Fischer, Y., Bauer, A., Beyerer, J.: A Conceptual Framework for Automatic Situation Assessment. In: *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, USA* (2011)
13. Shimada, Y., Hamaguchi, T., Takeda, K., Kitajima, T., Aoyama, A., Fuchino, T.: Study on Safety Operation Support System by Using the Risk Management Information. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) *KES 2006. LNCS (LNAI)*, vol. 4252, pp. 553–560. Springer, Heidelberg (2006)
14. Kalantarnia, M., Khan, F., Hawboldt, K.: Dynamic risk assessment using failure assessment and Bayesian theory. *Loss Prevention in the Process Industries* 22, 600–606 (2009)
15. Ardi, S.: Failure detection on safety components in a polyvinyl chloride batch process. *European Journal of Scientific Research* 1, 46–58 (2011)
16. U.S. Chemical Safety and Hazard Investigation Board: Vinyl Chloride Monomer Explosion. *Investigation Report* (2004)
17. Niu, L., Lu, J., Zhang, G.: *Cognition-Driven Decision Support for Business Intelligence-Models, Techniques, Systems and Applications*. Springer, Heidelberg (2009)

On Partial Comparability and Fuzzy Preference-Aversion Models

Camilo Franco, Javier Montero, and J. Tinguaro Rodríguez

Faculty of Mathematics, University Complutense, 28040 Madrid, Spain
francodelosrios@gmail.com, {monty, jtrodri}@mat.ucm.es

Abstract. A general overview of partial comparability and preference theory allows examining the notion of bipolarity and its role in the development of some general preference structures. This bipolar approach comes natural to the framework of decision theory, where different preference structures can be initially explored according to the type of bipolar model that they follow. Therefore, we compare two general preference structures, the first one, referred to as the PCT structure, which results from a well known axiomatic model for partial comparability theory, and the second one, referred to as the P-A structure, which extends one particular standard fuzzy preference model, such that some basic differences as well as particular similarities are clearly identified.

Keywords: Preference-aversion, fuzzy preference structures, partial comparability.

1 Introduction

Recently the notion of bipolarity has been used to make reference to the direct identification of opposites (see e.g., [2], [12]), allowing a dialectic construction of knowledge from reality. Hence, the nature of knowledge and its meaning rests upon the continuous search for distinct and antagonistic reference poles, where the positive information regarding one of the poles is taken as negative regarding the other one.

In this paper we take a quick view over the genealogy of decision theory, examining first the partial comparability theory and its outranking relation [10], [13], where bipolarity is naturally present as the principle of concordance-discordance. In the next section, standard fuzzy preference structures are reviewed [3], [9], identifying a certain bipolar interpretation. Then, in Section 4, a preference-aversion model is introduced, generalizing one standard fuzzy preference structure (a proposal initially presented in [4]), and finally, in Section 5, we compare it with a continuous extension of a model of partial comparability [11], [18], [19].

2 Partial Comparability

From a historical perspective, preference theory generally assumed the existence of a complete order over the set of alternatives. For example, in classical decision theory, where such alternatives were referred to as abstract utilities [21] or acts [14], it was

assumed that the individual either preferred or not one option over another. Otherwise, by exclusion, options were defined as indifferent or mathematically equivalent. But what seems most relevant there, from our present and practical point of view, is that lack of preference not necessarily makes reference to indifference, but also to incomparability, ignorance or in general, some explicit (negative) conflict on how to resolve a decision problem.

Beyond classical decision theory, some of its foundational postulates stand as the starting point from where a more general and flexible but equally consistent theory can be obtained. In this line, the inclusion of incomparability in the basic preference structure is an example of one theory which generalizes the classical complete order, by means of a mathematical structure that allows the construction of a partial order (see e.g., [1], [13]).

Consider for example the definition of a complete order such that for every x, y in X either $P(x,y)$ or $P(y,x)$ holds, being P a transitive preference relation that could be read as “being better than” (see e.g., [21]). In any other case, if none of them hold, then both alternatives are indifferent or equivalent, i.e., $x=y$. Now, this definition could be generalized for a strict partial order, if we say that for every x, y in X , at most one of the following relations hold: $x=y$, $P(x,y)$ or $P(y,x)$, being P transitive. In this way, any two alternatives in X can be incomparable, if none of those relations hold.

This idea is present in the initial proposal of the partial comparability axiom (PC), a first general approach to partial comparability theory (see [13]). There, it is assumed that the binary comparison between the alternatives of a given set X generates four basic relations, composing a complete preference structure given by the basic relations of indifference I (symmetric and reflexive), strict preference P (asymmetric and irreflexive), weak preference R (given that there is no possible distinguishability between P and I) and incomparability J (symmetric and irreflexive).

Besides, PC theory treats intransitive indifference and incomparability as basic preference relations and not as irrational or inconsistent ones. Under this approach [13], two methods for representing individual preferences are defined, the first one refers to a *foundational relational system of preference*, where only one of the basic relations (P , R , I and J) is assigned to any pair of alternatives, and the second one defines a *combined relational system of preference*, where a combination of the basic relations is assigned.

Focusing on the second method, i.e., on a *combined relational system of preference*, it is possible to identify a situation where no clear positive reasons exist for deciding if one alternative x is strictly preferred or indifferent to another alternative y , or, on the contrary, if y is strictly preferred to x . This situation is represented by the construction of a fuzzy *outranking* relation $OR(x,y)$ [10], [13].

In this way, given a certain criterium g over X , where $g(x)$ represents a positive subjective evaluation of x , the discrimination strength between the basic situations of preference can be examined by the *outranking* relation defined as $OR(x,y)=g(x)-g(y)$ [13]. Here, following the concordance-discordance principle, the values of $g(x)$ and $g(y)$ are interpreted as two antagonistic expressions, where the first one represents the positive or concordant attributes and the second one the negative or discordant attributes for preferring x , when it is compared with y .

Given the fuzzy character of this *outranking* preference approach [10], we now examine one standard fuzzy preference model, examining the more general structure

(P, I, J) . Notice that, although the PC theory recognizes the existence of the basic relation J , it is not included in the *outranking* framework: if $OR(x,y)$ is positive and no significant difference exists between x and y , i.e., if for a certain threshold r , $OR(x,y)$ is not greater than $r(g(y))$, then they are indifferent; if instead $OR(x)$ is positive and significant, i.e., if for a certain threshold p , $OR(x,y)$ is greater than $p(g(y))$, then there exists strict preference of x over y ; and if $OR(x,y)$ stands between $r(g(y))$ and $p(g(y))$, then there is weak preference between x and y .

3 Fuzzy Preference Relations

Standard preference models (see e.g., [3], [9]) assume a complete decomposition of individual preferences, such that the preference intensity between any pair of alternatives can be separated into strict preference P , indifference I , and incomparability J . This approach follows the classical preference structure for R , which in set theoretic form can be expressed as, $P = R \cap R^d$, $I = R \cap R^{-1}$, $J = R^c \cap R^d$, where R^c is the complementary of R , such that $R^c(x, y) = 1 - R(x, y)$, R^{-1} is the inverse of R , defined as $R^{-1}(x, y) = R(y, x)$, and R^d is the dual of R , such that $R^d = (R^{-1})^c$.

Then, such preference structure $R=(P, I, J)$ follows the three properties,

$$P \cup I = R \quad , \tag{1}$$

$$P \cup I \cup P^{-1} = R \cup R^{-1} \quad , \tag{2}$$

$$P \cup J = R^d \quad . \tag{3}$$

On the other hand (following [5], [17], [22]), a fuzzy preference binary relation is characterized by the fuzzy set $R(x, y) = \{ \langle x, y, \mu_R(x, y) \rangle \mid x, y \in X \}$, where $\mu_R(x, y): X \times X \rightarrow [0,1]$ is the membership function of R and $\mu_R(x, y) \in [0,1]$ is the membership intensity for every $x, y \in X$, according to the verification of the property of being “at least as good as”.

Here we follow one standard fuzzy preference model [3], [9], which defines fuzzy relations P, I, J in terms of the fuzzy preference relation R , using fuzzy logical operations that preserve as much as possible (1)-(3). This model assumes three general axioms: Independence of Irrelevant Alternatives (IA): for any pair of alternatives $x, y \in X$, the values of $P(x,y)$, $I(x,y)$ and $J(x,y)$ depend only on the values of $R(x,y)$ and $R(y,x)$, and are given by the functions $p, i, j: [0,1]^2 \rightarrow [0,1]$; Positive Association (PA): the function $p(x', y')$ is non-decreasing regarding its first argument and non-increasing regarding its second argument (assuming positive association on p is sufficient for assuring the same on i and j), where $x' = R(x, y)$ and $y' = R(y, x)$; and Symmetry (SM): i and j are symmetric (for a more detailed exposition see [3], [9]).

Now, it is necessary to specify the fuzzy logic operators for representing the valued disjunction and conjunction between fuzzy sets. Hence, these operators, represented for the case of the disjunction by a continuous t -conorm S , and for the conjunction by a continuous t -norm T , along with a negation n , should verify the De Morgan law [3], [15], $n(S(x', y')) = T(n(x'), n(y'))$, such that the three classical properties (1)-(3) can be reformulated as follows,

$$S(p(x', y'), i(x', y')) = x' \quad (4)$$

$$S(p(x', y'), i(x', y'), p(y', x')) = S(x', y') \quad (5)$$

$$S(p(y', x'), j(y', x')) = n(x') \quad (6)$$

Taking such operators as the strong Lukasiewicz-type De Morgan triple $\langle T, S, n \rangle_\varphi$, defined up to a certain automorphism φ of the unit interval, $\langle p, i, j \rangle_\varphi$ is a solution of (4)-(6) if and only if [3],

$$p(x', y') = \min\{x', n(y')\} \quad (7)$$

$$i(x', y') = T(x', y') \quad (8)$$

$$j(x', y') = T(n(x'), n(y')) \quad (9)$$

where i and j are mutually exclusive.

This preference model can be explored now as in the PC's outranking relation, where y' can be interpreted as the existing discordant information against x' . Notice that this approach includes incomparability as a basic relation (6), where a simultaneous verification of the strict or weak preference intensities (representing concordant information in favor of x) and their inverses and negations (acting as discordant information against x), can be verified (like in (7)-(9)).

4 Fuzzy Preference Aversion Relations

Different studies in social sciences have shown that losses loom larger than gains in decision problems under uncertainty [7], [20]. More generally stated, in the face of decision, individuals value distinctively reasons in favor or against the available alternatives [16]. This kind of gains-losses rationality allows the separate identification of certain harmful situations from possible desirable ones, representing in a more flexible and complete way the distinct epistemic states of a rational individual.

Following [2] (but also [12]), the aggregation of positive and negative affects can be characterized under a typology for *bipolarity*. The first type uses a univariate scale to map the individual's perceptions (see e.g., [6], [7], [20]) while a second type uses a bivariate scale, where such perceptions can be at the same time contradictory (as it

has been seen in Section 3). Under this framework, a preference-aversion model (P-A) is introduced where an *independent* aggregation of gains and losses is possible [4].

As input data, four intensities or membership degrees are defined by the functions $\mu_{R^+}(x, y), \mu_{R^+}^{-1}(x, y): X \times X \rightarrow [0,1]$ and $\mu_{R^-}(x, y), \mu_{R^-}^{-1}(x, y): X \times X \rightarrow [0,1]$, where such relations represent the strength with which any pair of elements x and y in X verify the two predicates R^+ and R^- . In this way, the fuzzy relation $R^+(x, y) = \left\{ \langle x, y, \mu_{R^+}(x, y) \rangle \mid x, y \in X \right\}$ represents the predicate “ x is at least as *good* as y ”, such that $\mu_{R^+}(x, y) \in [0,1]$, and $R^-(x, y) = \left\{ \langle x, y, \mu_{R^-}(x, y) \rangle \mid x, y \in X \right\}$ represents the predicate “ x is at least as *bad* as y ”, where $\mu_{R^-}(x, y) \in [0,1]$.

The inclusion of these two weak relations and their inverses allows the characterization of six binary relations, three for the positive preference order, which are the already known strict preference P , indifference I and incomparability on weak preference J , and the other three for the negative aversion order, which are strict aversion Z , negative indifference G , and incomparability on weak aversion H , such that the basic preference structure $\langle R^+, R^- \rangle = \langle (P, I, J), (Z, G, H) \rangle$ can be formally characterized by extending the three general axioms IA-SM as in IA(2)-SM(2) [4].

The axiom IA(2) states that for any two alternatives $x, y \in X$, the values of $P(x,y)$, $I(x,y)$ and $J(x,y)$, depend only on the two intensities $\mu_{R^+}(x, y)$ and $\mu_{R^+}(y, x)$, while the values of $Z(x,y)$, $G(x,y)$ and $H(x,y)$ depend only on $\mu_{R^-}(x, y)$ and $\mu_{R^-}(y, x)$, given by six continuous functions $p, i, j, z, g, h: [0,1]^2 \rightarrow [0,1]$; the axiom PA(2) affirms that the functions $p(x^+, y^+)$, $z(x^-, y^-)$ are non-decreasing regarding their first argument and non-increasing regarding their second argument, $i(x^+, y^+)$, $g(x^-, y^-)$ are non-decreasing regarding both arguments, and $j(x^+, y^+)$, $h(x^-, y^-)$ are non-increasing regarding both arguments, where $x^+ = \mu_{R^+}(x, y)$, $y^+ = \mu_{R^+}(y, x)$, $x^- = \mu_{R^-}(x, y)$ and $y^- = \mu_{R^-}(y, x)$; and the axiom SM(2) states the symmetry of $i(x^+, y^+)$, $j(x^+, y^+)$, $g(x^-, y^-)$, and $h(x^-, y^-)$.

Hence, as done in (4)-(6), where $x' = x^+$ and $y' = y^+$, we have the following properties for the negative order,

$$S(z(x^-, y^-), g(x^-, y^-)) = x^- \quad , \tag{10}$$

$$S(z(x^-, y^-), g(x^-, y^-), z(y^-, x^-)) = S(x^-, y^-) \quad , \tag{11}$$

$$S(z(y^-, x^-), h(y^-, x^-)) = n(x^-) \quad . \tag{12}$$

Like in the case of the preference order $\langle p, i, j \rangle_\phi$, $\langle z, g, h \rangle_\phi$ fulfills (10)-(12) if and only if its solution is given by,

$$z(x^-, y^-) = \min\{x^-, n(y^-)\} , \tag{13}$$

$$g(x^-, y^-) = T(x^-, y^-) , \tag{14}$$

$$h(x^-, y^-) = T(n(x^-), n(y^-)) . \tag{15}$$

where g and h are mutually exclusive.

Therefore, the complete system of equations, (4)-(6) and (10)-(12), that frames alternatives in terms of gains and losses type of rationality, allows to construct one positive order, based on the four relations $\langle p, p^{-1}, i, j \rangle_\phi$, and a negative order over X , based on the four relations $\langle z, z^{-1}, g, h \rangle_\phi$. In this way, the antagonism over the weak preference x^+ , is verified by the direct evaluation of its negative attributes, given by x^- , where $n(x^+) \neq x^- \neq y^+$, as opposed the PC approach and the outranking relation OR , or to the standard preference IA-SM model, where $n(x^+) \neq x^- = y^+$.

Now there are four independent subjective measures, two for the positive perception on gains, x^+ and y^+ , and two for the negative perception on losses, x^- and y^- . Hence, following the IA(2)-SM(2) model, the combination of the preference and aversion values, representing the different epistemic states of the individual, is characterized by ten distinct compound relations, which are specified in Table 1. The preference-aversion (P-A) structure, ($PZ, PA, PG, PH, IZ, IG, IH, JZ, JG$ and JH), is defined by the minimum operator between the basic preference and aversion relations.

Examining the P-A structure, those ten compound relations are obtained in the form of a system that represents the gradable situations that arise from the independent reasoning over gains and losses. In this way, a complete valuation space is defined, like a flexible framework for mapping the subjective perceptions of a rational individual in the face of decision.

Such valuation space is composed by ten different compound relations: (1) *incomparability due to ambivalence* (PZ), where two antagonistic valuations for strict preference and strict aversion coexist, (2) *strong strict preference* (PA), given by the intersection of strict preference and the inverse of strict aversion, such that strict preference for x becomes even stronger given that y is worse, (3) *pseudo-strict preference* (PG), a situation in which strict preference exists although the two alternatives are equally bad, (4) *semi-strict preference* (PH), where strict preference is verified but no information on the losses regarding both options is found.

Examining positive indifference and its combination with strict aversion, we obtain (5) *pseudo-strict aversion* (IZ), where the two alternatives are equally good but one is worse than the other, (6) *strong indifference* (IG), when both alternatives are equally good and bad, (7) *semi-indifference* (IH), where they are equally good but there is no available information on their negative attributes.

Continuing with incomparability due to ignorance on weak preference and the basic relations of the weak aversion dimension, we have (8) *semi-strict aversion* (JZ), with no verification of positive attributes under the presence of strict aversion, (9) *negative semi-indifference* (JG), with ignorance on the positive attributes of both options, which at the same time are equally bad, and finally, (10) *incomparability due to ignorance* (JH), reflecting absolute ignorance.

Table 1. The ten compound preference-aversion relations

$R = R^+, R^-$	$Z(x, y)$	$Z(y, x)$	$G(x, y)$	$H(x, y)$
$P(x, y)$	$PZ(x, y)$	$PA(x, y)$	$PG(x, y)$	$PH(x, y)$
$P(y, x)$	$PA(y, x)$	$PZ(y, x)$	$PG(y, x)$	$PH(y, x)$
$I(x, y)$	$IZ(x, y)$	$IZ(y, x)$	$IG(x, y)$	$IH(x, y)$
$J(x, y)$	$JZ(x, y)$	$JZ(y, x)$	$JG(x, y)$	$JH(x, y)$

We take into account that this is not the only way in which positive and negative arguments can be aggregated. An alternative model has been developed in the literature of preference theory, in particular, one axiomatic model for the PC theory [11], [18], [19], where ten basic relations are defined but under a *dependent* methodology. Such methodology evaluates first the coexistence of positive and negative reasons and then, from the resulting (four-valued) extension of each relation and its inverse (following the axiom of *independence of irrelevant alternatives* and by means on the *DDT* language [18]), constructs the corresponding preference PCT structure. In the following section we examine this approach and examine the relation between both structures (P-A and PCT).

5 The Four Valued Continuous Model of Partial Comparability

The four valued continuous extension for the PC theory [11], takes the intensities of weak preference and non-preference $\{x^+, x^-\}$ apart from the ones of $\{y^+, y^-\}$ and obtains an overall four valued weak preference predicate, along with its respective inverse, based on the four epistemic states given by [11],

$$R^i(x, y) = \min\{x^+, n(x^-)\} \quad , \quad (16)$$

$$R^f(x, y) = \min\{n(x^+), x^-\} \quad , \quad (17)$$

$$R^u(x, y) = T(n(x^+), n(x^-)) \quad , \quad (18)$$

$$R^k(x, y) = T(x^+, x^-) \quad . \quad (19)$$

So, if the weak preference and non-preference extensions are combined with their inverses, the following PCT structure is obtained [18], [19]: (1) strict preference, p ; (2) weak preference, q ; (3) semi preference, k ; semi weak preference, l ; strict indifference, i ; semi indifference, j ; weak indifference, h ; incomparability, r ; semi incomparability, u ; and weak incomparability, v (see Table 2 for the complete PCT structure).

Table 2. The ten basic preference situations of the PCT theory

PCT	$R^i(y, x)$	$R^k(y, x)$	$R^u(y, x)$	$R^f(y, x)$
$R^i(x, y)$	$i(x, y)$	$h(x, y)$	$k(x, y)$	$p(x, y)$
$R^k(x, y)$	$h(y, x)$	$j(x, y)$	$l(x, y)$	$q(x, y)$
$R^u(x, y)$	$k(y, x)$	$l(y, x)$	$u(x, y)$	$v(x, y)$
$R^f(x, y)$	$p(y, x)$	$q(y, x)$	$v(y, x)$	$r(x, y)$

As a result, ten basic preference relations are defined, the same number than in the P-A structure, but if we examine them in detail, we identify that their different methodology sets them apart both in their syntax and in their semantics, although they relate in a very similar way if a certain condition is forced.

Theorem: The PCT structure $(p, q, k, l, i, j, h, r, u, v)$, and the P-A structure $(PZ, PA, PG, PH, IZ, IG, IH, JZ, JG, JH)$, are only equivalent if we assume that $y^+ = x^-$.

Proof: Take the intersection of the basic P-A relations of (Z, Z^{-1}, G, H) and (P, P^{-1}, I, J) and of the PCT extensions of (R^i, R^k, R^u, R^f) and their inverses, as given by the continuous t -norm $\wedge = \min$. Also take the solutions for the basic P-A relations as given by equations (7)-(9) for preference and (13)-(15) for aversion, and take the PCT extensions as given by (16)-(19).

(1) PZ is not equivalent to i , as it follows from the definition of $PZ = \min(x^+, n(y^+)) \wedge \min(x^-, n(y^-))$ and $i = \min(x^+, n(x^-)) \wedge \min(y^+, n(y^-))$.

But if we assume $y^+ = x^-$, then they are trivially equivalent.

(2) IZ is not equivalent to h^{-1} , as it follows from the definition of $IZ = T(x^+, y^+) \wedge \min(x^-, n(y^-))$ and $h^{-1} = \min(y^+, n(y^-)) \wedge T(x^+, x^-)$. But, in the same way as before, if we assume $y^+ = x^-$, then they are equivalent.

(3) JZ is not equivalent to j^{-1} , as it follows from the definition of $JZ = T(n(x^+), n(y^+)) \wedge \min(x^-, n(y^-))$, $j^{-1} = \min(y^+, n(y^-)) \wedge T(n(x^+), n(x^-))$.

Again, if $y^+ = x^-$, then they are the same.

Hence, it can be verified that the following relations are not equivalent: PZ^{-1} and p , PG and h , PH and j , $P^{-1}Z$ and p^{-1} , $P^{-1}Z^{-1}$ and r , $P^{-1}G$ and q^{-1} , $P^{-1}H$ and v^{-1} , IZ^{-1} and q , IG and k , IH and l , JZ^{-1} and v , JG and l^{-1} , JH and u , unless it is assumed that $y^+ = x^-$. \diamond

In this way, the P-A and PCT structures are equivalent if and only if $y^+ = x^-$, case in which both proposals collapse into the same functional form of the IA-SM model, such that

$$R^i(x^+, x^-) = \min\{x^+, n(x^-)\} = \min\{x^+, n(y^+)\} = p(x^+, y^+) \quad (20)$$

$$R^f(x^+, x^-) = \min\{n(x^+), x^-\} = \min\{n(x^+), y^+\} = p(y^+, x^+) \quad (21)$$

$$R^u(x^+, x^-) = T(n(x^+), n(x^-)) = T(n(x^+), n(y^+)) = j(x^+, y^+) \quad (22)$$

$$R^k(x^+, x^-) = T(x^+, x^-) = T(x^+, y^+) = i(x^+, y^+) \quad (23)$$

In fact, if the condition $y^+ = x^-$ is forced, then the available information reduces, and by (20)-(23), the IA-SM approach is sufficient for its complete interpretation. But in the case in which $y^+ \neq x^-$, the PCT and the P-A structures specify the corresponding valuation map for the individual that orders the alternatives in a *dependent* (PCT case) or an *independent* (P-A case) manner, according to his/her perception on gains and losses and how they ought to be aggregated. Then, given the semantic space of the gains and losses rationality, the final outcome will differ according to its interpretation and the chosen methodology.

6 Conclusion

Modeling preference from two main more or less opposite forces is a standard approach in decision making, quite often establishing differences by stressing the meaning of each component of the model. Following [8], the next step should be to develop a comparative study of our approach with respect to other existing approaches, focusing on the underlying structural properties.

Acknowledgment. This research has been partially supported by the Government of Spain, grant TIN2009-07901.

References

1. Aumann, R.: Utility theory without the completeness axiom. *Econometrica* 30, 445–462 (1962)
2. Dubois, D., Prade, H.: An introduction to bipolar representations of information and preference. *International Journal of Intelligent Systems* 23, 866–877 (2008)

3. Fodor, J., Roubens, M.: *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer Academic Publishers, Dordrecht (1994)
4. Franco, C., Montero, J., Rodríguez, J.T.: Aggregation weights for a preference-aversion model. In: *Proceedings World Conference on Soft Computing*, San Francisco (2010), paper 199
5. Goguen, J.: The logic of inexact concepts. *Synthese* 19, 325–373 (1969)
6. Grabisch, M., Labreuche, C.: A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *A Quarterly Journal of Operations Research* 6, 1–44 (2008)
7. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291 (1979)
8. Montero, J., Gómez, D., Bustince, H.: On the relevance of some families of fuzzy sets. *Fuzzy Sets and Systems* 158, 2429–2442 (2007)
9. Montero, J., Tejada, J., Cutello, C.: A general model for deriving preference structures from data. *European Journal of Operational Research* 98, 98–110 (1997)
10. Perny, P., Roy, B.: The use of fuzzy outranking relations in preference modelling. *Fuzzy Sets and Systems* 49, 33–53 (1992)
11. Perny, P., Tsoukiàs, A.: On the continuous extension of a four valued logic for preference modeling. In: *Proceedings IPMU Conference*, Paris, France, July 6–10, pp. 302–309 (1998)
12. Rodríguez, J.T., Franco, C., Montero, J.: On the relationship between bipolarity and fuzziness. In: *Proceedings EUROFUSE Conference*, Régua, Portugal, September 21–23 (2011)
13. Roy, B.: Partial preference analysis and decision-aid: the fuzzy outranking relation concept. In: Bell, D., Keeney, R., Raiffa, H. (eds.) *Conflicting Objectives in Decisions*. Wiley and Sons, New York (1977)
14. Savage, L.: *The Foundations of Statistics*. Dover Publications, New York (1972)
15. Schweizer, B., Sklar, A.: *Probabilistic Metric Spaces*. North-Holland, Amsterdam (1983)
16. Smith, A.: *The Theory of Moral Sentiments*. Cambridge University Press, Cambridge (2002)
17. Trillas, E.: On a model for the meaning of predicates (A naïve approach to the genesis of fuzzy sets). In: Seising, R. (ed.) *Views of Fuzzy Sets and Systems from Different Perspectives*, pp. 175–205. Springer (2009)
18. Tsoukiàs, A., Vincke, P.: A new axiomatic foundation of partial comparability. *Theory and Decision* 39, 79–114 (1995)
19. Tsoukiàs, A., Vincke, P.: Extended preference structures in MultiCriteria Decision Aid. In: *Multicriteria Analysis*. Springer, Berlin (1997)
20. Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297–323 (1992)
21. Von Neumann, J., Morgenstern, O.: *Theory of Games and Economic Behavior*. Princeton University Press, Princeton (1953)
22. Zadeh, L.: Fuzzy sets. *Information and Control* 8, 338–353 (1965)

Honesty-Rate Measurement: A Novel Approach for the Fragile Trust Inside the DIDS

Peijian Chen¹, Yuexiang Yang², Hailong Wang¹, Chuan Tang², and Jie He¹

¹ School of Computer Science, National University of Defense Technology,
Changsha, 410008, China

² Information Center, National University of Defense Technology, Changsha, 410073, China
{chenpeijian, yyx, hlwang}@nudt.edu.cn, tangchuan@gmail.com,
jack.237@163.com

Abstract. In this paper, a novel honesty-rate measuring based approach is proposed to improve the security and trust of distributed intrusion detection systems. All the cooperative nodes join the system with an initial value of 1 for an honesty rate. The honesty rate of a node dynamically varies depending on its status and the honesty rate will decrease when a suspicious behavior of the node is detected and increase when the node obeys certain criteria. The proposed approach compares the honesty rate of each node to eliminate or reduce the impact of harmful information from the malicious nodes, and then reduces the false positives and false negatives of the intrusion detection systems. The experiments and analyses of a representative case confirm the ability of the proposed approach to improve detection accuracy and detection capability.

Keywords: Honesty-rate Measuring, Distributed, Intrusion Detection, OSSEC.

1 Introduction

In recent years, attack mechanisms on the Internet have been increasingly sophisticated, especially those that targeting large scale networks. Currently [1], cyber attackers scan vast scale of hosts from different IP domains for system or software vulnerabilities. They can spread computer worms, Trojan horses, or backdoors to infect and control a large number of hosts to establish botnets. And these botnets could be further exploited to launch Distributed Denial-of-Service (DDoS) [2] attacks to disrupt the service of a targeted network or system by overloading its capacity. These types of attacks are referred as large-scale coordinated attacks. They pose an extremely difficult challenge to the existing common Intrusion Detection Systems (IDSs). In order to detect these types of attacks, the IDSs need to integrate and analyze the evidence of suspicious network activities and system logs from multiple network domains. So that Distributed Intrusion Detection Systems (DIDSs) are needed to detect and response to large-scale coordinated attacks.

However, most existing DIDSs are built on the ground of trustness to the target systems and networks; they don't take much account of problems of the security and trust for DIDSs themselves. INDRA [3], which is considered to be the best approach

towards intrusion detection, uses a trusted network and generates alarms inside the network to provide information about suspected hosts. The greatest weak point of this approach is that it needs trust within the internal network as intrusions may rise from anywhere even inside the trusted network. Although it uses a neighborhood watching mechanism to make sure that alerts come from a trusted participant, it can't prevent legitimate users from sending malicious messages, which in turn will pollute the data analyzed to detect intrusion. DOMINO [4] aims to monitor large-scale coordinated attacks. This approach uses central certificate authority to validate the sources of messages, but it is still unable to deal with malicious messages which will drive the whole system in an unstable state. Cai Bo-qing (2008) proposed a node-linkage algorithm [5] based on information provided by snort, which uses collaborative communication between nodes to share the rules. Xue Yandong [6] introduced the subjective trust theory and feedback theory into cooperative detection, which reduces the probability of misjudge, and improves self-adapting capability. But it will be a disaster to the whole system when there are a few nodes more powerful than others, for they will be mistaken for malicious nodes.

2 Honesty-Rate Measuring Method

Before describing the whole system, we define the terminologies used in the Honesty-rate measuring system as follows:

- T-rate: Threshold honesty rate, below which a node can't be selected as a monitor node.
- N_x : Unique identification of each node.
- L : A numerical severity level ($0 \sim 15$) for behavior and status of a node classification in the view of the whole system.
- $H(N_x)$: Honesty-rate for node N_x , range from 0 to 1.
- $L(N_x)$: Severity level for N_x in the view of the whole system.
- L_x^y : Severity level for N_y in the view of node N_x .
- $S(N_x)$: A set of nodes whose H-rate are larger than T-rate and $H(N_x)$.

In the proposed system, all nodes are assumed to be honest [7] and join the network with an initial value of 1 for H-rate. The H-rate of a node dynamically increases or decreases depending on its behavior and status. The H-rate of a node is penalized when it performs some suspicious action like scanning other nodes. In contrast, a node is rewarded by increasing its H-rate when it behaves well, but its H-rate will not be larger than 1. The H-rate of a node is recomputed based on its H-rate, and the severity level in the views of the whole system and the monitor nodes. The workflow chart of honesty-rate measuring system is shown in Fig. 1.

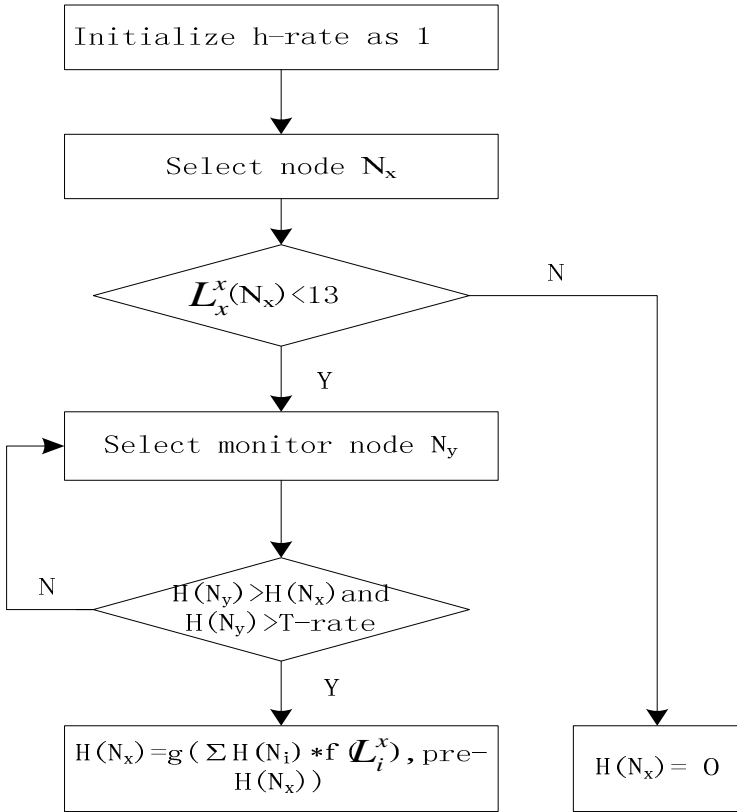


Fig. 1. Workflow chart for honesty-rate measuring system

2.1 Severity Level

The system has a numerical severity level ranging from 0 to 15, with 0 being the lowest and 15 the highest just like OSSEC [8]. The severity level of a node is based on degree of access attempts like:

- User generated events (access attempts) - level 5
- High importance events (unauthorized use) - level 12
- Severe attacks- level 15

The severity level can be changed on a case-by-case basic but the default severity levels are defined as follows in Table 1.

Table 1. Severity Level

Level	Description
0	Ignored, no action taken
1	None
2	System low priority notification
3	Successful/authorized events
4	System low priority errors
5	User-generated errors
6	Low relevance attacks
7	Bad words matching
8	First time seen
9	Error from invalid source
10	Multiple user generated errors
11	Integrity checking warning
12	High-importance event
13	Unusual error (high importance)
14	High importance security event.
15	Attack successful. Very small chance of false positive. Immediate attention is necessary.

The severity levels are more detailed in this system than others in order to enhance the precision of the attack classification. The severity level L is projected to a range from 0 to 1, in the function f , when it is used to compute the H-rate. Function f is used to adjust the size of L to avoid the change of L bringing a strong jitter in H-rate.

2.2 Monitor Nodes

The monitor nodes for N_x are selected from $S(N_x)$. For example, if $N_y \in S(N_x)$, then $H(N_y) > H(N_x)$ && $H(N_y) > H - rate$. In theory, the monitor nodes should include all the nodes in the set of $S(N_x)$. But too many monitor nodes will overload the system and take too much time. In practice, we balance the trade-off between the detection accuracy and the system performance. We usually pick up n nodes as monitor nodes from $S(N_x)$. The system can be more accurate with a larger n and there would be more system overhead as well. A node is randomly selected as a monitor node from $S(N_x)$, which makes the task to influence the monitor nodes even more difficult.

The monitor node has a set of data L_x^y [9], which contains the severity level of the other nodes in its view. For example, the monitor node N_x maintains a set of data $L_x^0 L_x^1 L_x^2 \dots L_x^i$. The number L is used to measure the behavior of other nodes. The set of data depends on the detection results for the malicious activities of other nodes in the network. We assume that L_x^y depends on its own status severity level when x equals y .

2.3 Measurement of H-Rate

Each node, node N_x for example, in the network stores a data set in the form of $\{L_x^0 L_x^1 L_x^2 \dots L_x^i, H(N_x)\}$. Measurement of $H(N_x)$ is divided into two cases, one is based on its status, and the other is based on its behavior metric L in the views of its monitor nodes. Node N_x is suspected to be an unreliable node, if $L_x^x \geq 13$, which means that the node N_x found itself had been attacked. The $H(N_x)$ will be 0, as shown in formula (1). If $L_x^x < 13$, $H(N_x)$ will be recomputed as a function of its pre- $H(N_x)$ and the severity level in the view of its monitor node by formula (2).

$$H(N_x) = 0, L_x^x \geq 13 \tag{1}$$

$$H(N_x) = g\left(\sum_1^n H(N_i) * f(L_i^x), pr - H(N_x)\right), L_x^x < 13 \tag{2}$$

2.4 Honesty-Rate Measuring Based DIDS

The workflow of the system based on honesty-rate measuring system is shown in Fig. 2. Firstly, the information collected will be sent to the filter module to filter the data of each node according to the filtering rules. This can effectively remove the data of which H-rate is lower than T-rate. Secondly, based on the honesty-rate measuring system and the filtered data, the DIDS revises the H-rate of each node and the filtering rules. Finally, the system places weight on the response rule according to H-rate of each node before the match and then generate an alert or record in the database.

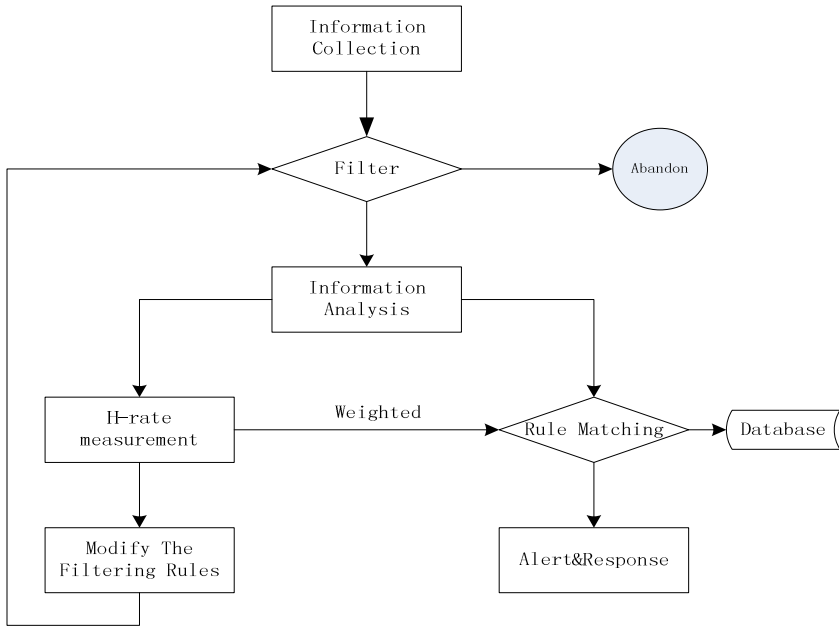


Fig. 2. Workflow chart of the DIDS

3 Experiments and Results

In this paper, the honesty-rate measuring system is introduced into the typical distributed intrusion detection system OSSEC. A new intrusion detection system TBO (Trust based OSSEC) is proposed. Fig. 3 shows the deployment of TBO. There are three types of clients in the TBO system: One is the OSSEC agent combined with snort, one is the OSSEC agent combined with the software firewall, and the other is the network equipment just like the hardware firework or router and so on. All the information of the host and the network behavior is sent to server for analysis in the form of log through UDP port 1514. Nodes like switches and routers are closed systems and thus are rather difficult to break in. So that their H-rates are assumed to be 1, always involve the measure of the H-rate of other nodes.

We have a simple verification for the TBO system in the laboratory LAN.

3.1 Experimental Settings

The system Includes one server (ubuntuServer10.10, 192.168.56.1), one attacking host (WindowsXp, 192.168.56.2, referred to be Unit 2), and seven collaboration hosts (one host WindowsXp, 192.168.56.3, referred to be Unit 3 and six hosts ubuntuDesktop10.10, the address range is 192.168.56.4 ~ 192.168.56.9).

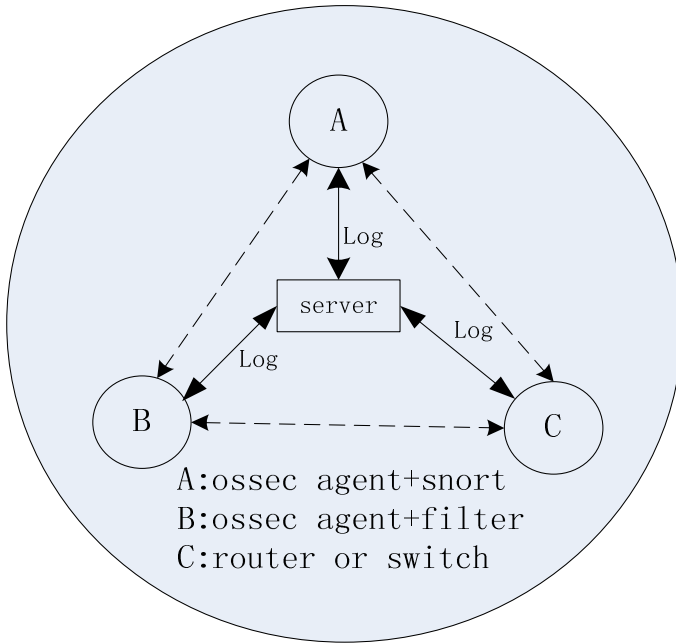


Fig. 3. The TBO system

3.2 Experimental Methods

The attacking host Unit 2 reports to the server, claiming that it is being attacked by Unit 7, which will pollute the data of the server and decrease the H-rate of Unit 3. There are two cases of experimental conditions, $L_{Unit2}^{Unit2} < 13$ and $L_{Unit2}^{Unit2} \geq 13$, which means that the Unit 2 reports its status to be reliable or unreliable respectively. The Threshold T-rate is set to be 0.7, and the number of monitor nodes n is set to be 4, according to the experimental conditions. For the limit of the OSSEC log analysis convergence speed, the time interval for TBO detection is 1 minute.

3.3 Results and Analysis

Fig. 4 shows the results of the experiment. When the attacking host Unit 2 reports itself to be unreliable ($L_{Unit2}^{Unit2} \geq 13$), all its reports to the server will not be accepted and will not influence the H-rate of Unit 3 as shown in Fig. 4 Series 1. And the TBO system will not alert or response to Unit 3. When $L_{Unit2}^{Unit2} < 13$, the malicious reports from Unit 2 claiming that being attacked by Unit 3 will influence the H-rate of Unit 3. As shown in Series 2, The H-rate of Unit 3 will jitter, and the jitter amplitude obeys a Poission distribution as the monitor was selected randomly. When $L_{Unit2}^{Unit2} < 13$ and Unit 2 has other malicious behavior which will decrease its H-rate, as is shown in Series 3, the TBO isolates Unit 3 only in the time of 4 minutes. As the H-rate of Unit

2 decreases, the influence of Unit 2 for Unit 3 will decrease, and then the H-rate of Unit 3 will increase. Before employing the honesty-rate measuring system, OSSEC would isolate Unit 3 whenever Unit 2 claimed to be attacked by Unit 3.

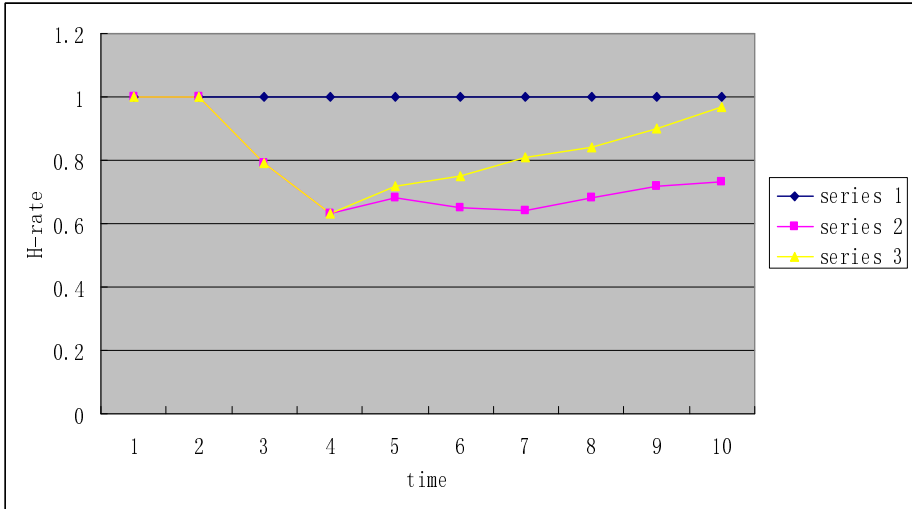


Fig. 4. The variation of the H-rate

4 Conclusion

To improve the security and trust for distributed intrusion detection systems (DIDSs), a new honesty-rate measuring based approach has been introduced to eliminate or reduce the impact of harmful information from the malicious nodes in this paper. A new intrusion detection system TBO has been built to test and verify the new honesty-rate measuring system can effectively improve detection accuracy and enhance detection capability. In terms of the future work, we will deploy the system in real setting, and evaluate its performance on real network traffic.

References

1. Zhou, C.V., Leckie, C., Karunasekera, S.: A survey of coordinated attacks and collaborative intrusion detection. *Computers & Security*, 124–140 (2010)
2. Zhang, C., Yin, J., Cai, Z., Chen, W.: RRED: robust RED algorithm to counter low-rate denial-of-service attacks. *Communications Letters* 14, 489–491 (2010)
3. Janakiraman, R., Waldvogel, M., Zhang, Q.: Indra: A peer-to-peer approach to network intrusion detection and prevention (2003)
4. Yegneswaran, V., Barford, P., Jha, S.: Global intrusion detection in the domino overlay system. In: *Proceedings of Network and Distributed Security Symposium, NDSS* (2004)
5. Cai, B.Q.: *Distributed Intrusion Detection System node cooperation algorithm*. Nanjing University of Science (2008)

6. Xue, Y.D., Han, X.L., Dong, S.F.: Distributed Cooperative Intrusion Detection System Based on Snort. *Computer Engineering* 36, 165–167 (2010)
7. Sen, P., Chaki, N., Chaki, R.: HIDS: Honesty-Rate Based Collaborative Intrusion Detection System for Mobile Ad-Hoc Networks. In: *7th Computer Information Systems and Industrial Management Applications*, pp. 121–126 (2008)
8. Bray, R., Cid, D., Hay, A., Corporation, E.: *OSSEC host-based intrusion detection guide*. Elsevier (2008)
9. Dou, W., Wang, H.M., Jin, Y., Zhou, P.: A recommendation-based peer-to-peer trust model. *Journal of Software* 4 (2004)

Using Cooperative Clustering to Solve Multiclass Problems

Chuanhuan Yin¹, Shaomin Mu², and Shengfeng Tian¹

¹ School of Computer and Information Technology, Beijing Jiaotong University,
Beijing, 100044, China

² School of Computer and Information Engineering, Shandong Agriculture University,
Taian, 271018, China

chyin@bjtu.edu.cn, msm@sdau.edu.cn, sftian@bjtu.edu.cn

Abstract. In this paper, we present a multiclass classification algorithm to address the multiclass problems with cooperative clustering. Using cooperative clustering, the cluster centers of all classes can be computed iteratively and simultaneously. In the process of clustering, we select a pair of adjacent class, and make their cluster center drawn towards the boundary. Therefore, the data set around a class is found and the data set plus the data in this class can be trained to form a classifier. With this algorithm, training efficiency and classification efficiency are improved with a slight impact on classification accuracy.

Keywords: Multiclass Classification, One-VS-All, Support Vector Machine, Cooperative Clustering.

1 Introduction

Many classification algorithms are specifically designed for binary problems, such as Support Vector Machines (SVMs) [1]. However, real world applications require that inputs are mapped into several categories. Building a multiclass classifier in one step needs to solve a larger optimization problem which is computationally more expensive. Several methods have been proposed to solve multiclass problems by combining several binary classifiers [2]. Let $S = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$ be training data set, where \mathbf{x}_i belongs to a domain $X \subset R^d$ and y_i is an integer from the set $Y = \{1, \dots, M\}$. The task is to find the function $f: X \rightarrow Y$ that maps an instance \mathbf{x} onto an element of Y . We can make a brief introduction of the proposed methods which approach multiclass problems as many simpler two-class problems.

1) One-vs-All (OVA). This method constructs M binary classifiers, each of which is trained against the rest $M - 1$ classes [3]. When an example is classified, M times binary tests are required to make final decision with majority voting.

2) One-vs-One (OVO). This method constructs $M(M - 1)/2$ binary classifiers, each of which separates a pair of classes [4]. When an example is classified, $M(M - 1)/2$ times binary tests are required with simple voting scheme. Another approach of combining the obtained classifiers is Directed Acyclic Graph (DAG) [5]. This method

constructs a rooted binary acyclic graph using binary classifiers. In the test phase, the number of binary tests is reduced to $M - 1$.

3) Error Correcting Output Codes (ECOC). In this method, Error Correcting Output Codes are used to improve the decision accuracy [6]. According to the encoding scheme, the number of binary tests range from M to $M(M - 1)/2$.

4) Binary Tree Architecture (BTA). In this method, a hierarchical binary decision tree is used to improve the training and testing efficiency [7], [8]. BTA only needs to train $M - 1$ binary classifiers and test $\log_2 M$ times for the classification decision.

The approach we proposed is based on a technique called cooperative clustering [9]. With this technique, the size of data set for training every binary classifier can be reduced to the size of two classes approximately in OVA.

This paper is organized as follows. In Section 2, we describe the method of cooperative clustering. In the next section, we illustrate the multiclass classification with OVA strategy. In Section 4 we give the results of the proposed methods. Finally, the conclusion is given in Section 5.

2 Cooperative Clustering

K-means is a widely used unsupervised learning algorithm that solves the well known clustering problem [10, 11]. The task of clustering is: given a set of n data points, denoted by $X = \{x_i \mid x_i \in R^d, i = 1, \dots, n\}$ and the number of clusters, c , each x_i has to be assigned to a cluster label $l_i \in \{1, \dots, c\}$. The algorithm consists of a simple re-estimation procedure as follows. Initially, c centers are assigned, one for each cluster. In step one, every point is assigned to the cluster whose center is closest to that point. In step two, the centers are computed for each cluster. These two steps are alternated until a stop criterion is met, i.e., when there is no further change in the assignment of the data points.

The cooperative clustering procedure is originally used for training support vector machines [9]. We iteratively compute the cluster centers of two classes simultaneously and draw them towards the boundary of the two classes. Then these cluster centers near the boundary can be used to replace the support vectors approximately. Because the training examples near the boundaries between adjacent classes carry more classification information than the samples in the centers of classes, the cooperative clustering is a good method to select support vectors for reducing the training of support vector machines. In this paper, we use cooperative clustering procedure to select data points surrounding a given class. These data points plus those in the given class can be used to train the binary classifier in OVA approach. In this way, the data set for training binary classifier is balanced, whereas the imbalance of examples in various classes is a fatal drawback of OVA strategy.

Suppose that there are two data sets X^+ and X^- , which belong to class + and class - respectively. We run k-means algorithm in the two data sets separately. The data set X^+ is partitioned into p clusters $\{X_1^+, X_2^+, \dots, X_p^+\}$ and the data set X^- into q clusters $\{X_1^-, X_2^-, \dots, X_q^-\}$, where $p \leq q$. Let $\mathbf{v}^+ = \{\mathbf{v}_1^+, \dots, \mathbf{v}_p^+\}$ be the set of cluster centers in

class +, $\mathbf{v}^- = \{\mathbf{v}_1^-, \dots, \mathbf{v}_q^-\}$ be the set of cluster centers in class -, and H be distance matrix between two sets \mathbf{v}^+ and \mathbf{v}^- . The ij -th entry h_{ij} of matrix H can be computed as

$$h_{ij} = \|\mathbf{v}_i^+ - \mathbf{v}_j^-\|^2 \quad i = 1, \dots, p, j = 1, \dots, q \quad (1)$$

We take out each pair $(\mathbf{v}_a^+, \mathbf{v}_b^-)$ of cluster centers with smallest distance from sets \mathbf{v}^+ and \mathbf{v}^- iteratively according to matrix H . The two cluster centers in $(\mathbf{v}_a^+, \mathbf{v}_b^-)$ should move towards another. Let r_a^+ be the average radius of cluster a in class + and r_b^- be the average radius of cluster b in class -. We have

$$r_a^+ = \frac{1}{n_a^+} \sum_{\mathbf{x} \in X_a^+} \|\mathbf{x} - \mathbf{v}_a^+\| \quad (2)$$

$$r_b^- = \frac{1}{n_b^-} \sum_{\mathbf{x} \in X_b^-} \|\mathbf{x} - \mathbf{v}_b^-\| \quad (3)$$

where n_a^+ and n_b^- are sample numbers in clusters X_a^+ and X_b^- respectively. Then each pair $(\mathbf{v}_a^+, \mathbf{v}_b^-)$ is updated as follows

$$\mathbf{v}_a^+ = \mathbf{v}_a^+ + \lambda \frac{r_a^+}{r_a^+ + r_b^-} (\mathbf{v}_b^- - \mathbf{v}_a^+) \quad (4)$$

$$\mathbf{v}_b^- = \mathbf{v}_b^- + \lambda \frac{r_b^-}{r_a^+ + r_b^-} (\mathbf{v}_a^+ - \mathbf{v}_b^-) \quad (5)$$

where $\lambda \in (0, 1)$ is the quantity controlling the distance between \mathbf{v}_a^+ and \mathbf{v}_b^- . With the newly produced cluster centers, the two data sets X^+ and X^- are partitioned again. After several iterations, p pairs of cluster centers near the boundary are obtained.

The whole procedure of the cooperative clustering is as follows.

- (1) Partition X^+ into clusters $\{X_1^+, X_2^+, \dots, X_p^+\}$ and X^- into clusters $\{X_1^-, X_2^-, \dots, X_q^-\}$ with k-means.
- (2) Compute the cluster centers $\mathbf{v}^+ = \{\mathbf{v}_1^+, \dots, \mathbf{v}_p^+\}$ in class + and $\mathbf{v}^- = \{\mathbf{v}_1^-, \dots, \mathbf{v}_q^-\}$ in class -.
- (3) Set $V^s = \{\}$ and Compute the matrix $H = \{h_{ij}\}_{1 \leq i \leq p, 1 \leq j \leq q}$ with equation (1).
- (4) Find a smallest element h_{ab} in H .
- (5) Add $(\mathbf{v}_a^+, \mathbf{v}_b^-)$ into the set V^s .
- (6) Delete the row a and the column b from matrix H .
- (7) Update each center pair $(\mathbf{v}_a^+, \mathbf{v}_b^-)$ in set V^s with equations (4) and (5).
- (8) If the number of pairs in set V^s is less than p , go to step (4).

- (9) Partition X^+ into clusters $\{X_1^+, X_2^+, \dots, X_p^+\}$ and X^- into clusters $\{X_1^-, X_2^-, \dots, X_q^-\}$ with updated cluster centers.
- (10) If the partition is changed in step (9), go to Step (2).

With the above procedure, we can find p pairs of cluster centers. Each pair crosses the boundary of two classes. The time complexity of the procedure is (*#iterations*) $\times O(n(p + q))$.

3 Multiclass Classification with OVA Strategy

In this section, we adopt the OVA strategy and present a training method for binary classifier. Let $S = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$ be training data set, where \mathbf{x}_i belongs to a domain $X \subset R^d$ and y_i is an integer from the set $Y = \{1, \dots, M\}$. OVA method constructs M binary classifiers. The i th Classifier, f_i , is trained using all the samples of class i as positive instances and the samples of the other classes as negative instances. Usually the training data set is imbalanced. To obtain balanced training data set, we use cooperative clustering procedure as follows. Firstly, all the samples of class i are assigned as positive instances and the samples of the other classes as negative instances. Suppose that each class is partitioned into c clusters. Then there are c clusters in class + and $c(M - 1)$ clusters in class -. The cooperative clustering procedure is run on these samples. Then cM cluster centers are obtained in which $2c$ cluster centers are in pairs. The c pairs of cluster centers include c centers in class + and c centers in class -. These cluster centers correspond to two data sets: one includes the samples in class i , another includes the samples surrounding the class i . These two data sets are denoted as X^+ and X^- . Usually these two data sets are balanced and can be used to train classifier f_i . To improve the accuracy of the classifier, the cluster centers which are not in pairs are also included in dataset X^- .

The training procedure of the classifier f_i is described as follows.

- (1) Initialize sets $X^+ = \{\mathbf{x}_j \mid y_j = i, j = 1, \dots, n\}$ and $X^- = \{\mathbf{x}_j \mid y_j \neq i, j = 1, \dots, n\}$.
- (2) Set $p = c$ and $q = c(M - 1)$.
- (3) Run the cooperative clustering procedure. Obtain the partition $\{X_1^-, X_2^-, \dots, X_q^-\}$, the set of cluster centers $\mathbf{v}^- = \{\mathbf{v}_1^-, \dots, \mathbf{v}_q^-\}$, and the set of center pairs V^s .
- (4) For each $X_j^-, j = 1, \dots, q$, remove data points in X_j^- from X^- and add cluster center \mathbf{v}_j^- into X^- if corresponding center \mathbf{v}_j^- is not in the set V^s .
- (5) Training the classifier f_i with data sets X^+ and X^- .

With the above procedure, the size of data set for training a binary classifier is about the size of two classes and the numbers of positive and negative samples are approximately equal.

For instance, in Fig. 1, the data set of four classes is shown and the line denotes the boundary of two classes.

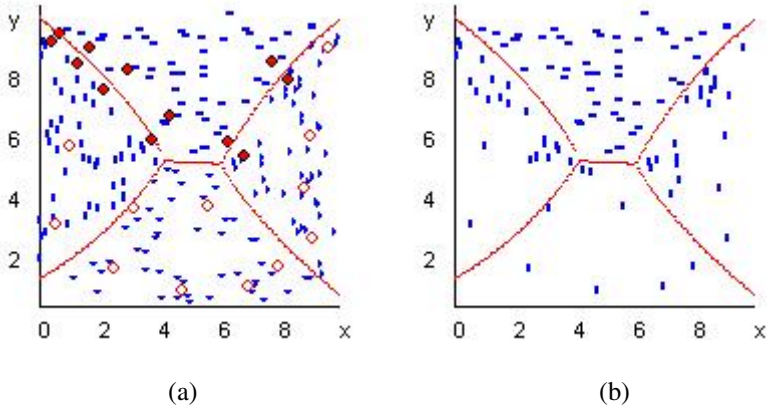


Fig. 1. A data set consists of four classes and the cooperative clustering procedure is performed for the top class. In (a), circles denote cluster centers and solid circles denote those in pairs. In (b), the data points for training the top class are shown with two different shapes.

In Fig. 1a, the cooperative clustering procedure is performed for the top class and the cluster centers are denoted by large circles. All of the six cluster centers in the top class are listed near its boundary. In other classes, only six cluster centers are constructed pairs with cluster centers in the top class. In Fig. 1b, only the training data are shown, which include the data in the top class and the data around the top class.

4 Experiments

To test the effectiveness of the proposed methods, we apply them to five benchmark data sets from the UCI Machine Learning Repository [12] summarized in Table 1. The experiments were done on a Pentium 4, 2.2 GHz with 256 MB RAM.

Table 1. Data sets used in the experiments

Problem	#training data	#testing data	#attribute	#class
vehicle	592	254	18	4
segment	1617	693	19	7
sat	4435	2000	36	6

In the experiments, we use support vector machine (SVM) as base learner [13]. There are two important parameters in SVM algorithm: C and σ , where C is a parameter that determines the tradeoff between the maximum margin and the minimum classification error, σ is a parameter used in RBF kernel $K(\mathbf{x}, \mathbf{x}') = \exp[-\|\mathbf{x} - \mathbf{x}'\|^2 / \sigma^2]$. Among many learning procedures for SVMs, we choose Platt's sequential minimal optimization (SMO) as training algorithm [14]. We fixed $C = 10$ and optimize σ in every application.

In cooperative clustering procedure, parameters λ and c are important for the performance of the classifier. In our experiments, $\lambda \in [0.6, 0.7]$ is good choice and we choose $\lambda = 0.6$ in the following experiments. The parameter c is used to control the accuracy and complexity. In the following experiments, c is set to about 10% of average size of training data in each class. For example, the problem sat has 4435 training data and six classes. We set $c = 4435/6/10 \approx 74$.

For problems vehicle and segment, we randomly choose 70% for training and test the remaining 30%. The results are the average values among the experiments. For problem sat, training sets and testing sets are provided separately in the UCI repository. For all problems, we linearly scale all training data to be in $[-1, 1]$ for features with plus and minus values and in $[0, 1]$ for features with only plus values. And the testing data are scaled accordingly.

We compare the classification performance with the following methods:

- (1) OVA [3]: ordinary one-vs-all method;
- (2) CC_OVA: one-vs-all method based on cooperative clustering described in Section 3;

We compare the classification performance with the above three methods in following seven aspects:

- (1) C, σ, c : parameters of SVM and cooperative clustering;
- (2) Training acc: the classification accuracy on training data;
- (3) #SVs: the average number of support vectors in each binary classifier;
- (4) #Training data: the number of training samples in every binary classifier;
- (5) Training time: the time for training all binary classifiers;
- (6) Testing acc: the classification accuracy on testing data;
- (7) Testing time: the time for classifying the testing data;

The experiment results for the five data sets are shown in the Table 2. Although the testing accuracy of CC_OVA is little lower than OVA in some cases, but other performances of CC_OVA are better than that of OVA: the size of training data and the training time decrease. The number of support vectors, i.e. the complexity of SVM, is reduced. As a result, the testing time decreases.

Table 2. The comparison of CC_OVA and OVA on three datasets

Parameter	Vehicle		Segment		Sat	
	OVA	CC_OVA	OVA	CC_OVA	OVA	CC_OVA
C, σ, c	10,0.5,-	10,0.5,14	10,0.5,-	10,0.5,23	10,0.2,-	10,0.2,74
Training acc(%)	91.05	88.51	98.89	98.76	100.0	99.98
#SVs	207	171	162	126	1737	956
#Training data	592	325	1617	586	4435	1876
Training time(s)	5.09	2.48	28.86	8.02	139.34	72.88
Testing acc(%)	77.17	77.17	97.55	97.26	91.95	91.95
Testing time(s)	0.11	0.06	1.56	0.55	12.47	5.20

5 Conclusion

In this paper, we have proposed a new method to improve the one-vs-all approach in multiclass classification using a new clustering method called cooperative clustering. In the cooperative clustering procedure, we iteratively compute the cluster centers of two classes simultaneously and draw them towards the boundary of the two classes. In this way, the cluster centers near the boundary are found. The data points in these clusters are used to train the binary classifier. With the above procedure, the size of data set for training a binary classifier is about the size of two classes and the numbers of positive and negative samples are approximately equal. Experiments show that the one-vs-all strategy for multiclass classification based on cooperative clustering is efficient and has a slight impact on generalization performance. In our future work, we plan to use the cooperative clustering procedure in local kernel machine on very large datasets.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (No. 61105056), the Fundamental Research Funds for the Central Universities, the Science and Technology Foundation of Beijing Jiaotong University (No. 2007RC066) and the youth Science and Technology Innovation Foundation of Shandong Agriculture University (No. 200923647).

References

1. Vapnik, V.N.: *Statistical Learning Theory*. John Wiley and Sons, New York (1998)
2. Hsu, C.-W., Lin, C.-J.: A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Networks* 13, 415–425 (2002)
3. Anand, R., Mehrotra, K.G., Mohan, C.K., Ranka, S.: Efficient classification for multiclass problems using modular neural networks. *IEEE Trans. Neural Networks* 6, 117–124 (1995)
4. Hastie, T., Tibshirani, R.: Classification by pairwise coupling. *The Annals of Statistics* 26, 451–471 (1998)
5. Platt, J., Cristianini, N., Shawe-Taylor, J.: Large margin DAGSVM's for multiclass classification. In: *Proceedings of Advances in Neural Information Processing System* (2000)
6. Dietterich, T.G., Bakiri, G.: Solving multiclass learning problems via error-correcting output codes. *J. Artificial Intelligence Research* 2, 263–286 (1995)
7. Cheong, S., Oh, S.H., Lee, S.-Y.: Support vector machines with binary tree architecture for multi-class classification. *Neural Info. Process.-Lett. Rev.* 2, 47–51 (2004)
8. Fei, B., Liu, J.: Binary tree of SVM: a new fast multiclass training and classification algorithm. *IEEE Trans. Neural Networks* 17, 696–704 (2006)
9. Tian, S., Mu, S., Yin, C.: Cooperative clustering for training SVMs. In: *Proceedings of 3th Int. Symposium on Neural Networks* (2006)
10. MacQueen, J.B.: Some Methods for classification and Analysis of Multivariate Observations. In: *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability* (1967)

11. Maulik, U., Bandyopadhyay, S.: Performance evaluation of some clustering algorithms and validity indices. *IEEE Trans. Pattern Recognition and Machine Intelligence* 24, 1650–1654 (2002)
12. Blake, C.L., Merz, C.J.: UCI repository of machine learning databases. Dept. Inform. Comput. Sci., Univ. California, Irvine, CA (1998)
13. Cristianini, N., Shawe-Taylor, J.: *An Introduction to Support Vector machines*. Cambridge University Press, Cambridge (2000)
14. Platt, J.C.: Sequential minimal optimization: a fast algorithm for training support vector machines. Microsoft Research, Tech. Rep. MSR-TR-98-14 (1998)

Total Colorings of Planar Graphs with Maximum Degree Seven and without 3-Cycles Adjacent to 5-Cycles

Guangde Liu¹, Bing Wang^{1,2}, and Jian-liang Wu²

¹ Department of Mathematics, Zaozhuang University, Zaozhuang, 277160, China

² School of Mathematics, Shandong University, Jinan, 250100, China

lgd@uzz.edu.cn, wangbing456@uzz.edu.cn, jlwu@sdu.edu.cn

Abstract. Let G be a planar graph with maximum degree $\Delta \geq 7$ and without 3-cycles adjacent to 4-cycles, that is, any 3-cycle has not a common edge with a 5-cycle. Then the total chromatic number of G is $\Delta + 1$.

Keywords: Total coloring, Planar graph, Cycle.

1 Introduction

All graphs considered in this paper are simple, finite and undirected, and we follow [2] for the terminologies and notations not defined here. Let G be a graph. We use $V(G), E(G), \Delta(G)$ and $\delta(G)$ (or simply V, E, Δ and δ) to denote the vertex set, the edge set, the maximum degree and the minimum degree of G , respectively. For a vertex $v \in V$, let $N(v)$ denote the set of vertices adjacent to v , and let $d(v) = |N(v)|$ denote the degree of v . A k -vertex, k^+ -vertex or a k^- -vertex is a vertex of degree k , at least k or at most k respectively. A k -cycle is a cycle of length k , and a 3-cycle is usually called a triangle.

A *total- k -coloring* of a graph G is a coloring of $V \cup E$ using k colors such that no two adjacent or incident elements receive the same color. The *total chromatic number* $\chi''(G)$ of G is the smallest integer k such that G has a total- k -coloring. Clearly, $\chi''(G) \geq \Delta + 1$. Behzad [1] and Vizing posed independently the following famous conjecture, which is known as the total coloring conjecture (TCC).

We will introduce some more notations and definitions here for convenience. Let $G = (V, E, F)$ be a planar graph, where F is the face set of G . The degree of a face f , denoted by $d(f)$, is the number of edges incident with it, where each cut-edge is counted twice. A k -face or a $+$ -face is a face of degree k or at least k , respectively. Denote by $n_d(v)$ the number of d -vertices adjacent to a vertex v .

Conjecture. For any graph G , $\chi''(G) \leq \Delta + 2$.

This conjecture was confirmed for a general graph with $\Delta \leq 5$. But for planar graph, the only open case is $\Delta = 6$ (see [2,10]). Interestingly, planar graphs with high maximum degree allow a stronger assertion, that is, every planar graph with high maximum degree Δ is $(\Delta + 1)$ -totally-colorable. This result was first established in [2] for $\Delta \geq 14$, which was extended to $\Delta \geq 9$ (see [9]). For $4 \leq \Delta \leq 8$, it is not known whether that the assertion still holds true. It is challenging to prove or disprove that planar graphs of maximum degree $\Delta \in \{4, \dots, 8\}$ are $(\Delta + 1)$ -totally-colorable. Such a study has attracted a considerable amount of attention and some neat results on this topic are obtained. Recently, Du *et. al.*[8] proved that planar graphs with maximum degree 8 and without adjacent triangles are 9-totally-colorable. Liu *et. al.*[11] proved that if G is a planar graph without intersecting 4-cycles and $\Delta \geq 7$, then $\chi''(G) = \Delta + 1$ and the authors of the paper proved the same result if G is a planar graph with $\Delta \geq 7$ and without intersecting 3-cycles. For every planar graph with girth at least 4, Bordin *et. al.*[4] proved that If $\Delta \geq 7$, then the planar graph is 8-totally-colorable. In [12], Shen and Wang proved the same result if planar graphs with $\Delta \geq 7$ and without 5-cycles. In this paper, we shall extend these results and prove the following result.

Theorem 1. *Let G be a planar graph without 3-cycles adjacent to 4-cycles. If $\Delta \geq 7$, then $\chi''(G) = \Delta + 1$.*

According to Theorem 1, we can draw a conclusion which generalizes related known result.

Corollary 2. [4,12] *Let G be a planar graph without k -cycles for some $k \in \{3, 5\}$. If $\Delta \geq 7$, then $\chi''(G) = \Delta + 1$.*

2 Proof of Theorem 1

Let G be a minimal counterexample to Theorem 1 in terms of the number of vertices and edges, that is, every proper subgraph of G is $(\Delta + 1)$ -totally-colorable, but G itself is not. So G is 2-connected and the boundary of each face in G is exactly a cycle (i.e. each face can not pass through a vertex v more than once). We first show some known properties on G .

(a) G contains no edge uv with $\min\{d(u), d(v)\} \leq 4$ and $d(u) + d(v) \leq 8$ (see [12]);

(b) The subgraph G_{27} of G induced by all edges joining 2-vertices to 7-vertices is a forest (see [3,5]);

(c) If a vertex v adjacent to two vertices v_1, v_2 such that

$$2 \leq d(v_1) = d(v_2) = 9 - d(v) \leq 3, \text{ then every face incident with } vv_1 \text{ or } vv_2$$

is a 4^+ -face (see [11]);

- (d) G contains no 3-faces incident with more than one 4-vertex (see [11]);
- (e) If v is a 7-vertex of G with $n_2(v) \geq 1$, then $n_{4^+}(v) \geq 1$ (see [7]).

(f) If a 7-vertex v is adjacent to three 2-vertices v_1, v_2, v_3 such that v, v_1, v_2 are incident with the same face f , then f is a 6^+ -face (see [8]).

Lemma 3. [7] Suppose v is a 7-vertex and v_1, v_2, \dots, v_k are consecutive neighbors of v with $d(v_1) = d(v_k) = 2$ and $d(v_i) \geq 3$ for $2 \leq i \leq k-1$, where $k \in \{3, 4, 5, 6\}$. If the face incident with v, v_i, v_{i+1} is a 4-face $vv_i x_i v_{i+1}$ for $1 \leq i \leq k-1$, then at least one vertex in $\{v_2, v_3, \dots, v_{k-1}\}$ is a 4^+ -vertex.

Since G is a planar graph, by Euler's formula, we have

$$\sum_{v \in V} (2d(v) - 6) + \sum_{f \in F} (d(f) - 6) = -12 < 0.$$

Now we define the initial charge function $ch(x)$ of $x \in V \cup F$ to be $ch(v) = 2d(v) - 6$ if $v \in V$ and $ch(f) = d(f) - 6$ if $f \in F$. It follows that

$$\sum_{x \in V \cup F} ch(x) < 0.$$

Now we design appropriate discharging rules and redistribute weights accordingly. Note that any discharging procedure preserves the total charge of G . If we can define suitable discharging rules to charge the initial charge function ch to the final charge function ch' on $V \cup F$, such that $ch'(x) \geq 0$ for all $x \in V \cup F$, then we get an obvious contradiction.

Our discharging rules are as follows.

R1. Every 2-vertex receives $3/2$ from its child and $1/2$ from its parent.

R2. Let f be a 3-face. If f is incident with a 3^- -vertex, then it gets $3/2$ from each of its incident 6^+ -vertices.

If f is incident with a 4^+ -vertex, then it gets 1 from the 4^+ -vertex and gets 1 from each of its incident 5^+ -vertices.

R3. Let f be a 4-face. If f is incident with two 3^- -vertices, then it gets 1 from each of its two incident 6^+ -vertices. If f is incident with only one 3^- -vertex, then it gets $1/2$ from the incident 4^+ -vertex and gets $3/4$ from each of its two incident 6^+ -vertices. If f is not incident with any 3^- -vertex, then it gets $1/2$ from each of its incident vertices.

R4. Every 5-face gets $1/3$ from each of its incident 4^+ -vertices.

The rest of this paper is to check $ch'(x) \geq 0$ for all $x \in V \cup F$. Firstly note that our discharging rules are just designed such that $ch'(f) \geq 0$ for all $f \in F$ and $ch'(v) \geq 0$ for all 2-vertices $v \in V$. So we only check that $ch'(v) \geq 0$ for all 3^+ -vertices G .

Let v be a vertex of G . If $d(v) = 3$, then

$ch'(v) = ch(v) = 0$. Suppose $d(v) = 4$. Then $t \leq 2$. If $1 \leq t \leq 2$, then v is incident with at least two 6^+ -faces and v sends at most 1 to each of its incident 3-faces by **R2**, and it follows that $ch'(v) \geq ch(v) - t \times 1 \geq 0$. Otherwise, v is incident with four 4^+ -faces. So $ch'(v) \geq ch(v) - \frac{1}{2} \times 4 = 0$

by **R3**. Suppose $d(v) = 5$. Then $t \leq 3$. Note that v sends at most 1 to each of its incident 3-face by **R2**, at most 1/2 to each of its incident 4^+ -faces by **R3**, at most 1/3 to each of its incident 5-faces by **R4**. If $1 \leq t \leq 3$, then v is incident with at least two 6^+ -faces, and it follows that

$$ch'(v) \geq ch(v) - 1 \times t - (5 - 2 - t) \times \frac{1}{2} = \frac{5 - t}{2} > 0. \text{ Otherwise, } v \text{ is incident}$$

with five 4^+ -faces. So $ch'(v) \geq ch(v) - \frac{1}{2} \times 5 > 0$. Suppose $d(v) = 6$. Then

$t \leq 4$. Note that v sends at most 3/2 to each of its incident 3-face by **R2**, at most 1 to each of its incident 4^+ -faces by **R3**, at most 1/3 to each of its incident 5-faces by **R4**. Thus, If $1 \leq t \leq 3$, by the same argument as above,

$$ch'(v) \geq ch(v) - \frac{3}{2} \times t - (6 - 2 - t) \times 1 = \frac{4 - t}{2} \geq 0.$$

Otherwise, $ch'(v) \geq ch(v) - 1 \times 6 \geq 0$.

Suppose $d(v) \geq 7$. Then $ch(v) = 2 \times 7 - 6 = 8$. If

$n_2(v) \geq 1$, then v sends at most $\frac{n_2(v) + 2}{2}$ to all its adjacent 2-vertices by

R1, at most 3/2 to each of its incident 3-face by **R2**, at most 1 to each of its incident 4^+ -faces by **R3**, at most 1/3 to each of its incident 5-faces by **R4**. Let v_1, v_2, \dots, v_7 be neighbors of v and f_1, f_2, \dots, f_7 be faces incident with v in an anticlockwise order, where f_i is incident with v_i and v_{i+1} , where $i \in \{1, 2, \dots, 7\}$. Note that all the subscripts in the paper are taken modulo 7.

Lemma 4. Suppose that $d(v_i) = d(v_k) = 2$ and $d(v_j) \geq 3$ for all $j = i+1, \dots, k-1$. If $f_i, f_{i+1}, \dots, f_{k-1}$ are 4^+ -faces, then v sends at most $3/2 + (k-3)$ (in total) to f_1, f_2, \dots, f_{k-1} .

Proof. By Lemma 3, $\max\{d(v_{i+1}), \dots, d(v_{k-1})\} \geq 4$ or $\max\{d(f_1), \dots, d(f_{k-1})\} \geq 5$. If $\max\{d(v_{i+1}), \dots, d(v_{k-1})\} \geq 4$, then v sends at most $2 \times 3/4 + (k-1-2)$ to f_i, \dots, f_{k-1} by **R2**. If $\max\{d(f_1), \dots, d(f_{k-1})\} \geq 5$, then v sends at most $1/3 + (k-1-1)$ to f_i, \dots, f_{k-1} by **R3** and **R4**. Since $2 \times 3/4 > 1 + 1/3$, v sends at most $3/2 + (k-3)$ to f_1, f_2, \dots, f_{k-1} .

By (e), we have $n_2(v) \leq 6$. So it suffices to consider the following cases.

Case 1. $n_2(v) = 6$. Then v is incident with at least five 6^+ -faces and not incident with any 3-face by (c) and (f). So $ch'(v) \geq ch(v) - (6+2)/2 - 3/2 > 0$ by Lemma 4.

Case 2. $n_2(v) = 5$. If v is not incident with any 3-face, that is, $t = 0$, then $ch'(v) \geq ch(v) - (5+2)/2 - 4 \times 1 > 0$. Otherwise, v is incident with at least four 6^+ -faces, and v sends at most $3/2$ to its incident 3-face f by **R2**. So $ch'(v) \geq ch(v) - (5+2)/2 - 3/2 - 2 \times 1 > 0$.

Case 3. $n_2(v) = 4$. There are four possible configurations as shown in figure 1

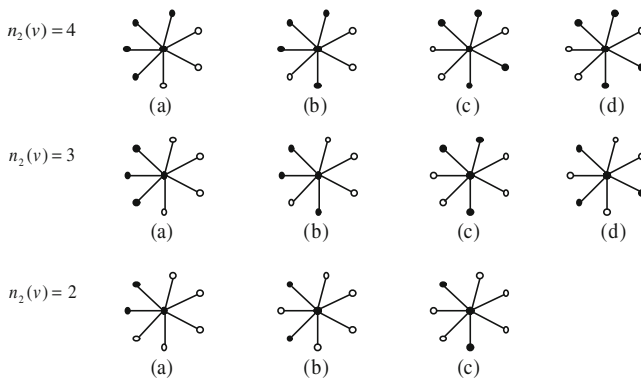


Fig. 1. Possible distributions of the adjacent 2-vertices of an 7-vertex with $n_2(v) \leq 4$

For figure 1, v is incident with at least one 6^+ -face and $t \leq 2$ by (c) and (f). If $t = 0$, then $ch'(v) \geq ch(v) - (4+2)/2 - \max\{3/2+2, 3/2+3/2+1, 3 \times 3/2\} > 0$ by Lemma 4. If $1 \leq t \leq 2$, then v is incident with at least another two 6^+ -faces, and it follows that

$$ch'(v) \geq ch(v) - \frac{4+2}{2} - \frac{3}{2} \times t - \frac{1}{3} \times 2 - (7-1-2-t) \geq 0 \text{ by our rules.}$$

Case 4. $n_2(v) = 3$. There are four possible configurations as shown in figure 1.

For figure 1, $t \leq 2$ by (c) and (f). If $t = 0$, then

$$ch'(v) \geq ch(v) - \frac{3+2}{2} - \max\left\{\frac{3}{2}+3, \frac{3}{2}+\left(\frac{3}{2}+2\right), 2 \times \left(\frac{3}{2}+1\right), 2 \times \frac{3}{2} + \left(\frac{3}{2}+1\right)\right\} > 0$$

by Lemma 4. If $t = 1$, then v is incident with at least two 6^+ -faces, and it follows that $ch'(v) \geq ch(v) - \frac{3+2}{2} - \frac{3}{2} - (7-2-1) = 0$

If $t = 2$, then v is incident with at least another 6^+ -faces f_6 , and it follows that

$$ch'(v) \geq ch(v) - (3+2)/2 - t \times 3/2 - (7-1-2-t) = (3-t)/2 > 0 \text{ by our rules.}$$

Case 5. $n_2(v) = 2$. There are three possible configurations as shown in figure 1.

For (a), we have that the face f_6 is a 5^+ -face, and it follows that f_6 receives at most $1/3$ from v by **R4**. So

$ch(v) - (2+2)/2 - 1/3 = 8 - 2 - 1/3 = 17/3$. If $1 \leq t \leq 3$, then v is incident with at least two 6^+ -faces, and it follows that

$$ch'(v) \geq \frac{17}{3} - t \times \frac{3}{2} - (7-1-2-t) = (10-3t)/6 > 0;$$

Otherwise, $ch'(v) \geq \frac{17}{3} - \left(\frac{3}{2}+4\right) > 0$ by Lemma [4].

For (b) and (c), if $t = 0$, then

$$ch'(v) \geq ch(v) - \frac{2+2}{2} - \max\left\{\frac{3}{2}+\frac{3}{2}+3, \left(\frac{3}{2}+1\right)+\left(\frac{3}{2}+2\right)\right\} > 0 \text{ by Lemma 4. If}$$

$1 \leq t \leq 2$, then v is incident with at least two 6^+ -faces, and it follows that

$$ch'(v) \geq ch(v) - \frac{2+2}{2} - \frac{3}{2} \times t - (7-2-t) = \frac{2-t}{2} \geq 0.$$

If $t = 3$, then v is incident with at least three 6^+ -faces, it follows that

$$ch'(v) \geq ch(v) - \frac{2+2}{2} - 3 \times \frac{3}{2} > 0$$

by our rules.

Case 6. $n_2(v) = 1$. Note that $n_{4^+}(v) \geq 1$ by (e) and $t \leq 4$. If $t = 0$, then

$$ch'(v) \geq ch(v) - \frac{1+2}{2} - 2 \times \frac{3}{4} - 5 = 0$$

or $ch'(v) \geq ch(v) - \frac{1+2}{2} - \frac{1}{3} - 6 > 0$ by Lemma 4 and **R4**. If $1 \leq t \leq 3$ and

v is incident with at least two 6^+ -faces, it follows that

$$ch'(v) \geq ch(v) - \frac{1+2}{2} - \frac{3}{2} \times t - (7-2-t) = \frac{3-t}{2} > 0. \text{ If } t = 4, \text{ then } v \text{ is inci-}$$

dent with at least 6^+ -faces, and it follows that

$$ch'(v) \geq ch(v) - \frac{1+2}{2} - 4 \times \frac{3}{2} > 0.$$

Case 7. $n_2(v) = 0$. Note that $t \leq 4$. If $1 \leq t \leq 4$, it follows that

$$ch'(v) \geq ch(v) - t \times \frac{3}{2} - (7-2-t) = \frac{6-t}{2} > 0. \text{ By our rule. Otherwise,}$$

$$ch'(v) \geq ch(v) - 7 \times 1 > 0$$

by **R3**. Hence we complete the proof of the Theorem 1.

References

1. Behzad, M.: Graphs and their chromatic numbers, Ph. D. Thesis, Michigan State University (1965)
2. Bondy, J.A., Murty, U.S.R.: Graph Theory with Applications. MacMillan, London (1976)
3. Borodin, O.V.: On the total coloring of planar graphs. J. Reine Angew. Math. 394, 180–185 (1989)
4. Borodin, O.V., Kostochka, A.V., Woodall, D.R.: List edge and list total colourings of multigraphs. J. Combin. Theory Ser. B 71, 184–204 (1997)
5. Borodin, O.V., Kostochka, A.V., Woodall, D.R.: Total colorings of planar graphs with large maximum degree. J. Graph Theory 26, 53–59 (1997)
6. Borodin, O.V., Kostochka, A.V., Woodall, D.R.: Total colourings of planar graphs with large girth. Europ. J. Combin. 19, 19–24 (1998)
7. Chang, G., Hou, J., Roussel, N.: Local condition for planar graphs of maximum degree 7 to be 8-totally-colorable (submitted)
8. Du, D., Shen, L., Wang, Y.: Planar graphs with maximum degree 8 and without adjacent triangles are 9-totally-colorable. Discrete Applied Math. 157, 6035–6043 (2009)

9. Kowalik, L., Sereni, J.-S., Skrekovski, R.: Total colorings of planar graphs with maximum degree nine. *SIAM J. Discrete Math.* 22, 1462–1479 (2008)
10. Kostochka, A.V.: The total chromatic number of any multigraph with maximum degree five is at most seven. *Discrete Math.* 162, 199–214 (1996)
11. Liu, B., Hou, J.F., Wu, J.L., Liu, G.Z.: Total colorings and list total colorings of planar graphs without intersecting 4-cycles. *Discrete Math* 309(20), 6035–6043 (2009)
12. Shen, L., Wang, Y.: Planar graphs with maximum degree 7 and without 5-cycles are 8-totally-colorable. *Discrete Mathematics* 310, 2372–2379 (2010)

A Meta-Strategy for Coordinating of One-to-Many Negotiation over Multiple Issues

Khalid Mansour and Ryszard Kowalczyk

Faculty of Information & Communication Technologies,
Swinburne University of Technology, Melbourne, Australia
{mwmansour, RKowalczyk}@groupwise.swin.edu.au

Abstract. This paper presents a novel approach for managing multi-bilateral concurrent negotiations. We extend our previous work by considering a situation where a buyer agent negotiates with multiple seller agents concurrently over multiple continuous issues instead of a single issue. A related work in this area considers a meta-strategy for bilateral negotiations. This work adapts the previous related work to coordinate multi-bilateral concurrent negotiations taking into consideration the different behaviors of the opponents during negotiation to decide on choosing the appropriate negotiation strategy (i.e., trade-off or concession) for the buyer agent's delegates at each negotiation round. A negotiation meta-strategy to coordinate the one-to-many negotiation form is proposed and empirically tested under various negotiation environments. The experiments show the robustness of our coordination mechanism.

Keywords: Negotiation, Coordination, Meta-Strategy.

1 Introduction

Negotiation is an important interaction mechanism for autonomous agents [9]. Automated agents use negotiation to resolve conflicts by exchanging offers and counteroffer to reach an agreement(s). During the past decade, automated negotiation domain witnessed a large momentum and a wide interest from researchers [1, 4, 8]. In this work, we consider the one-to-many negotiation form where an agent negotiates concurrently with many other agents. Many possible application domains can be represented by this form of negotiation, such as the supply chain domain, the task allocation and order fulfillment problems [18]. For example, a manufacturer needs to procure raw materials and in most cases, more than one supplier exists in the market. One of possible approaches for procuring raw materials is to negotiate with multiple potential suppliers concurrently. In the service oriented domain for example, a service requester can negotiate with multiple service providers concurrently over the SLAs for the purpose of procuring one service or more [11]. The advantage of concurrent negotiation is that an agent (i.e., the buyer agent in our case) can utilize the fact that many other agents compete for an agreement.

Previously we proposed a mechanism to change the buyer's negotiation strategy by changing the convexity degree of the buyer's concession curve [10]. In

this paper, we propose a meta-strategy to manage the concurrent negotiations conducted by a buyer agent with multiple independent seller agents over multiple continuous issues. Our coordination meta-strategy is based on the idea of alternating between two negotiation strategies, namely, the concession strategy (e.g., [4]) and the trade-off strategy (e.g., [6]). Both strategies are offer generation mechanisms. The main difference between them is that the trade-off strategy generates different counteroffers on the buyer's iso-curve (indifference curve) which have the same utility value of its previous negotiation round counteroffers' utility. On the other hand, the concession mechanism generates counteroffers with lower utility value (i.e., provide more value to the opponents) than the previous negotiation round counteroffers' utility.

Since the buyer agent negotiates with multiple sellers for the purpose of securing one agreement which the buyer agent will select at the end of negotiation, and to decrease the risk of not reaching an agreement, our approach classifies opponents at the beginning of negotiation and after few negotiation rounds into two groups, favorable group and unfavorable group depending on their concession behavior. The favorable group is the group that offers generous concession. For those favorable group, we apply our proposed meta-strategy and for the unfavorable group we use the concession strategy. More details can be found in section 4.

Having granted the privilege only for the buyer agent to renege from a deal without a penalty, while forcing the seller agents to honor their agreements can be a realistic scenario in situations where the number of seller agents is large and/or the seller agents are offering infinite supply (e.g. information.) In such cases, a seller agent might be satisfied to make deals with many potential buyers in a hope that some of these buyers will confirm their deals later.

The idea behind our meta strategy is to keep the buyer agent on the iso-curve of its current utility aspiration level for the longest possible time before moving to a lower utility iso-curve level. If an agent stays on its current iso-curve, it might reach an agreement with the same utility level of the current indifference curve. The downside of staying long time on the indifference curve and not conceding is the risk of not reaching an agreement [17]. After a certain number of negotiation rounds, an agent staying on its iso-curve can not keep proposing offers/counteroffer that may give the opponent better utility value (in case that agents are semi-cooperative and have divergent preferences over issues under negotiation.) The meta strategy should alternate between the trade-off and concession strategies in a way that increases the chance of reaching a valuable agreement for the buyer agent.

Our meta strategy is not only competitive, but also cooperative. It is competitive in the sense that it tries to achieve the best possible agreement for the buyer agent and cooperative in the sense that it tries to generate counteroffers that provide better values to the opponents assuming that the buyer agent and the seller agents have divergent preferences over issues, hence we propose an Iterative Offer Generation method (**IOG**) that can be competitive and cooperative at the same time.

This work adapts the work in [5] [17] by considering a more complex situation where one agent negotiates with multiple agents concurrently. The work in [5] [17] considers mainly a meta strategy in bilateral negotiation.

In our work, we assume that agents are competitive and self interested and negotiate under incomplete knowledge. In addition, the reservation values, negotiation deadlines and utility structures are private information for each agent.

The rest of the paper is organized as follows. Section 2 reviews the related work while section 3 is dedicated for the negotiation framework. Section 4 discusses our coordination approach. The experimental settings and evaluation is given in section 5. Finally, section section 6 concludes the paper and outlines the future work.

2 Related Work

During the last decade, work has been done to address the one-to-many form of negotiation as an alternative mechanism to the single-sided auction protocol [1] [2] [12] [14]. The first explicit architecture for the one-to-many negotiation form was presented in [16] where the buyer agent consists of sub-negotiators and a coordinator. The study proposes four different coordination strategies during concurrent multiple bilateral negotiation:(1) desperate strategy in which the buyer agent accepts the first agreement and quits negotiations with all other sellers (2) patient strategy where the buyer agent makes temporary agreements with seller agents during negotiation and holds on these agreements until all the remaining threads of negotiations are finished, then the buyer agent selects the agreement with the highest utility (3) optimized patient which is similar to the patient strategy except that it does not accept a new agreement with less utility than the highest existing one (4) finally, the manipulation strategies in which the coordinator changes the negotiation strategies of its sub-negotiators during negotiation. Our approach adopts a meta-strategy that alternates between a trade-off strategy and a concession strategy.

Several studies were published to address some aspects of the one-to-many negotiation form [10] [12] [16]. Most published works focus on the situation where agents negotiate over a single continuous issue (e.g., price) for the purpose of securing one agreement, while this paper investigates the problem of one-to-many negotiation over multiple continuous issues for the purpose of securing one agreement. Our work is similar to some existing work [14] [13] in terms of choosing the coordination approach that changes the negotiation strategy during negotiation. For example, a decision making technique for changing the negotiation strategies during negotiation depending on historic information of previous negotiations regarding of the agreements rate and the utilities rate is proposed in [14]. Our approach investigates the multi-issue negotiation domain and is based on the progress of the current negotiation encounter and does not rely on historic information.

While the work in [1] [12] considers decommitment penalty during negotiation, we assume that the buyer agent incurs no penalty for exercising decommitment during negotiation. Some heuristic methods were proposed to estimate

the expected utility in both a synchronized multi-threaded negotiations and a dynamic multi-threaded negotiations [3]. The synchronized multi-threaded negotiations model considers the existing outside options for each single thread, while the dynamic multi-threaded negotiations considers also the uncertain outside options that might arrive in the future. In both cases, the methods assume a knowledge of the probability distribution of the reservation prices of the opponents. In many cases, this kind of information is not available.

The above related work considers the one-to-many negotiation form where agents negotiate over a single issue, while our work investigates negotiations over multiple issues. A more relevant work to ours is the one presented in [5] [17] where the studies propose a meta-strategy for bilateral negotiation over multiple issues. The idea is that an agent chooses between using either a trade-off strategy or a concession strategy in each negotiation round. The agent uses a concession strategy when the distance between the last two offers of both agents was not reducing [5] or whenever it faces a deadlock situation in which the agent's opponent proposes an offer that has a lower utility value than the previous one [17], i.e., non monotonic concession behavior. The work in this paper differs from the works present in [5] [17] in a few aspects. First, this paper investigates the one-to-many form of negotiation which is a more complex setting than the bilateral form targeted in [5] [17]. Second, the non monotonic behavior of the opponent is not assumed. Another aspect is that we consider the number of outside options (which can not be considered as a factor in bilateral negotiation) and the different behavior of the opponents to decide on a certain offer generation strategy during negotiation.

Our work demonstrates the fact that an agent is able to take advantage of negotiating with multiple opponents concurrently over multiple continuous issues and provides a unique approach that guarantees better utility outcome for the agent.

3 Negotiation Framework

We consider a buyer agent \mathbf{b} negotiates with a set of seller agents $\mathbf{S} = \{s_1, s_2, \dots, s_n\}$ concurrently. The buyer agent has a set of delegate negotiators $\mathbf{D} = \{d_1, d_2, \dots, d_n\}$. Each delegate d_i negotiates with a seller s_i (see Figure 1) The buyer agent and the seller agents negotiate over a set of issues $\mathbf{J} = \{j_1, j_2, \dots, j_g\}$.

The buyer agent creates and destroys delegate negotiators during negotiation as a response to the number of the seller agents who enters or leaves negotiation. We assume that the seller agents are independent in their actions (i.e., they do not exchange information.) In our approach, the coordinator unit (see Figure 1) coordinates the actions of the buyer agent \mathbf{b} by deciding on the type of the offer generation strategy (e.g., conceding strategy) for each delegate $d_i \in \mathbf{D}$ in each negotiation round. Other possible coordination decisions can be related to the parameters associated with the specific offer generation technique (e.g., convexity degree of the concession curve ' β ' associated with the time-dependent tactics),

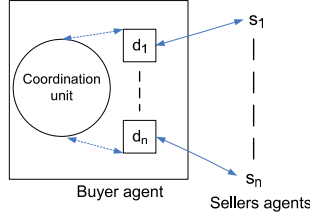


Fig. 1. One-to-Many Negotiation form

the reservation value etc. In this paper, we consider only selecting the type of offer generation strategy at each negotiation round, i.e., trade-off or concession.

We assume that agents use the alternate negotiation protocol [15] in which agents exchange offers and counteroffers at each negotiation round as an interaction protocol. Each agent has a negotiation deadline t_{max} by which the agent must accept an offer or withdraw from negotiation by that time. In addition, each agent has a reservation value for each issue. The reservation value of an issue is the minimum/maximum acceptable value for a certain issue during negotiation.

The utility (u) of a single issue j_i of value x at time t is calculated according to equation [1].

$$u_{j_i}^t = \begin{cases} (x_{j_i} - Min_{j_i}) / (Max_{j_i} - Min_{j_i}) & \text{If } IV_{j_i} > RV_{j_i} \\ (Max_{j_i} - x_{j_i}) / (max_{j_i} - Min_{j_i}) & \text{If } IV_{j_i} < RV_{j_i} \end{cases} \quad (1)$$

Where IV_{j_i} and RV_{j_i} refers to the agent’s initial value and the agent’s reservation value respectively. We assume that both **D** and **S** negotiate over the same issue set **J**. The utility of a given offer (an offer in our case consists of more than one issue value) is calculated as the weighted utility values of the issues in a given offer. Let U be the weighted utility of an offer X at time t , then

$$U_X^t = w_{j_1} u_{j_1}^t + w_{j_2} u_{j_2}^t + \dots + w_{j_g} u_{j_g}^t \quad (2)$$

Given that $\sum_{i=1}^g w_i = 1$. The weights of the issues ($w_{j_1}, w_{j_2}, \dots, w_{j_g}$) are given before the start of negotiation.

As an offer generation mechanism, we propose an iterative offer generation (**IOG**) method which is cooperative and competitive mechanism for generating offers. The (**IOG**) method works as follows, first an agent ranks the issues under negotiation according to their importance to the opponent in case of the agent uses a trade-off strategy and according to their importance to the agent in case of the agent uses a concession strategy. Second, if the agent chooses the trade-off strategy, it concedes first on most important issue from the opponent’s point of view, then Equations [1] and [2] are used to calculate the value of the least important issue (while using the previous values for all other issues) from the opponent’s point of view. That would actually propose a new value to the least

important issue that would keep the agent on its current iso-curve (i.e., $U_{X'}^{(t+1)} = U_X^t$.) If the value of the most important issue reaches its reservation value, then the agent starts conceding on the second important issue from the opponent's point of view and again uses the Equations 1 and 2 to find a new value for the second least important issue or the third, fourth, etc. least important issues that would keep the agent on its current iso-curve. The algorithm is repeated until the offer is ready. It is similarly used in case the agent uses a concession strategy except that all issues are assigned their initial values or little concession is offered (e.g., Boulware strategy 4) on each issue except for the least important issue from the agent's point of view which is calculated using Equations 1 and 2 given that a lower value for $U_{X'}^{(t+1)}$ is preassigned (i.e., $U_{X'}^{(t+1)} < U_X^t$). If the least important issue reaches its reservation value, then the agent starts conceding on the second lower important issue etc. The algorithm is repeated until the offer is ready.

At any point in time during negotiation, the buyer agent uses Equation 3 to decide on the minimum acceptable utility value for the next negotiation round.

$$U_X^t = (t/T^b)^{1/\beta^b} \quad (3)$$

Where T^b is the buyer's deadline and β^b is the buyer's concession parameter, i.e., the convexity degree of the buyer's concession curve.

For generating trade-off offers, the buyer agent detects for possible divergence in preferences with its opponents over issues under negotiation and uses that information to rank the issues accordingly and then generates counteroffers that provides each opponent with the highest possible value given that all counteroffers are on the current buyer's iso-curve. Whenever the buyer agent has divergent preferences with its opponents over certain issues, the buyer agent concedes on issues that provides higher value for its opponents more while staying on its indifference curve.

It is easy for an agent to discover whether it has divergent preferences with its opponent over certain issues by comparing the amount of concession on each issue from both sides (i.e., the agent and its opponent.) For example, if the issues under negotiation are price and quantity and the buyer agent offers \$10 and 50kg in the first negotiation round and \$30 and 48Kg in the second negotiation round, while the seller agent offers \$100 and 5Kg in the first round and \$90 and 11Kg in the second round. It is obvious that the seller is more sensitive towards the price when compared to the buyer while the buyer is more sensitive to the quantity when compared to the seller. In this case the buyer would concede more on the price (while staying on its indifference curve) while the seller may choose to concede more on quantity. The above example illustrates an easy method to detect for possible divergence in agents' preferences over issues under negotiation. Given that, when the buyer chooses to use the trade-off strategy, it detects first for possible divergence in preferences over issues and then generates its counteroffers accordingly using the **IOG** method.

At each negotiation round, the buyer agent needs to use either a concession strategy or a trade-off strategy, deciding on choosing the type of strategy and the

amount of concession if the strategy of choice is the concession one are discussed in the next section.

4 Meta-Strategy Coordination

The presented concurrent negotiation coordination technique involves change of negotiation strategy during negotiation. For example, in the proposed negotiation model (see section 2), the buyer agent may change the offer generation mechanism or change any of the parameters of a certain offer generation mechanism (e.g., the convexity degree of the concession curve associated with the time-dependent tactics [4].) Formally, let Ω_a be the negotiation strategy of an agent a , then $\Omega_a = \langle IV^a, RV^a, T^a, \Theta^a \rangle$, where $IV^a, RV^a, T^a, \Theta^a$ stand for the initial offer value, the reservation value, the deadline and the set of offer generation techniques of an agent a respectively. Our representation of an agent's strategy Ω_a is similar to [7], the difference is that the fourth part of the strategy in [7] represents the β value in the time dependent tactics [4] while the fourth part in our representation (Θ^a) has more general representation that indicates any offer generation method (e.g., trade-off, time-dependent, behavior dependent etc.)

Since our coordination model adopts a meta-strategy that uses either a trade-off strategy or a concession strategy in each negotiation round for delegates negotiating with the favorable group of the opponents, the buyer agent needs to decide on choosing a particular strategy for those delegates at the start of each negotiation round. Given that the negotiation form is one-to-many, the buyer agent can benefit from existing of multiple opponents to decide on choosing a certain strategy.

In general, the buyer agent uses the trade-off strategy whenever the situation is favorable, otherwise the buyer uses the concession strategy. Deciding whether the buyer agent is in a favorable situation depends on both, the number of opponents and their behaviors in terms of their concessions. If the opponents are offering generous concessions then the situation is favorable for the buyer agent, otherwise it is not. More opponents means more competitive advantage from the buyer's point of view.

Our approach works as follows, in the first few negotiation rounds the buyer agent classifies the opponents into two groups according to their relative behavior (relative to the buyer's behavior) by comparing the amount of concessions from each negotiation pair (i.e., (d_i, s_i) .) The buyer agent applies the meta-strategy on the favorable group, while uses the concession strategy with the unfavorable group.

At this point, we have two problems to deal with. The first one is how to measure an opponent's behavior? The second one is how much the buyer agent should concede when it chooses to use the concession strategy? The first problem can be solved by comparing the first order differences of offers/counteroffers from each negotiation pair (i.e., (d_i, s_i)) and count the number of seller agents (ϵ) having more concession than the concession offered by their counterparts of

the buyer's delegates for the same number of negotiation rounds (h). The buyer agent needs to select a threshold number $\bar{\epsilon}$ ($\bar{\epsilon} \geq \epsilon$) that is necessary for the buyer agent to choose the trade-off strategy, otherwise, it plays conceder.

Formally, let $\Delta f c_{s_i}^t = |(x_{s_i \rightarrow d_i}^t - x_{s_i \rightarrow d_i}^{t-h})|$ for both increasing and decreasing issue values of a seller s_i , $\Delta f c_{d_i}^t = |(x_{d_i \rightarrow s_i}^t - x_{d_i \rightarrow s_i}^{t-h})|$ for both increasing and decreasing issue values of a delegate d_i , where h stands for a number of steps back in the history of the current negotiation offers ($h \leq t$). At time t and for each (d_i, s_i) , we calculate the difference between $\Delta f c_{s_i}^t$ and $\Delta f c_{d_i}^t$ as shown in Equation 4

$$\forall (d_i, s_i), \Delta F C_i^t = \Delta f c_{s_i}^t - \Delta f c_{d_i}^t \quad (4)$$

Then add each $\Delta F C_i^t$ to a set $\Delta F C^t$. Now, count the number of elements in the set $\Delta F C^t$ having positive values (i.e., ϵ) and once we have ϵ and $\bar{\epsilon}$ then the buyer agent decides whether to use the trade-off strategy at a certain negotiation round.

The final problem which is selecting the appropriate amount of concession when the buyer agent chooses to use the concession strategy (for both groups, the favorable group and unfavorable one) can be managed by benefiting from existing of more than one opponent. One possible solution is to select the least amount of concession offered by an existing opponent and concedes similarly. Formally, let $\mathbf{C} = \{c_1, c_2, \dots, c_n\}$ where \mathbf{C} is the set of the first order differences of the last two negotiation round concessions in utility offered by the opponents, then the buyer agent chooses to concede with a value $\mathbf{v} = \text{Min}(\mathbf{C})$. This solution is competitive and reduces the risk of not reaching an agreement since there are still other agents who are conceding more than \mathbf{v} which increases the chance of reaching a valuable agreement to the buyer. Choosing \mathbf{v} as a value for the buyer's concession in the next negotiation round is a tit-for-tat variant in the one-to-many form of negotiation.

5 Experimental Settings and Results

We test our meta-strategy against large number of possible opponents types, the negotiation strategies of opponents combine both, time dependent tactics and behavior dependent tactics [4] as follows:

1. Time-dependent tactics:
 - 1- Boulware: $B = \{\beta | \beta \in [0.05, 0.7]\}$
 - 2- Linear: $L = \{\beta | \beta \in [0.9, 1.1]\}$
 - 3- Conceder: $C = \{\beta | \beta \in [2, 10]\}$, where β is the concession parameter.
2. Tit-For-Tat tactic:
 - Relative Ti-For-Tat, $\delta = 1$
3. The weights used to mix the time-dependent tactics and behavior-dependent tactics to generate an offer are divided in three groups
 - 1- Small, $S \in [0.05, 0.3]$
 - 2- Medium, $M \in [0.4, 0.6]$
 - 3- Large, $L \in [0.7, 0.9]$

In our experiments, we set $\bar{\epsilon} = 1, h = 1, |\mathbf{J}| = 2$. The buyer agent plays against different opponents types that cover infinite number of possible strategies:

$$St = \text{time_dependent tactics} \times \text{behavior_dependent tactic} \times \text{Weights}$$

$$St = (B \cup L \cup C) \times r \times (S \cup M \cup L).$$

Decomposing St results in the following set of 9 negotiation environments:

$$St = \{(B \times r \times S), (B \times r \times M), (B \times r \times L), (L \times r \times S), (L \times r \times M), (L \times r \times L), (C \times r \times S), (C \times r \times M), (C \times r \times L)\}.$$

The buyer agent plays against the St elements to test for the robustness of our algorithm. The overlap (the overlap between the buyer agent the seller agents reservation boundaries) percentages are selected randomly from the interval $[0.4,0.6]$.

When there are 2 seller agents, the total number of possible settings is 36 (using combination, $\binom{9}{2}$). For 4 opponents ($\binom{9}{4}$), the number of possible settings is 126 etc. For each setting, we run the experiment for 100 times. Which means that in the case of having 2 opponents, the total number experimental repetitions is 3600 times.

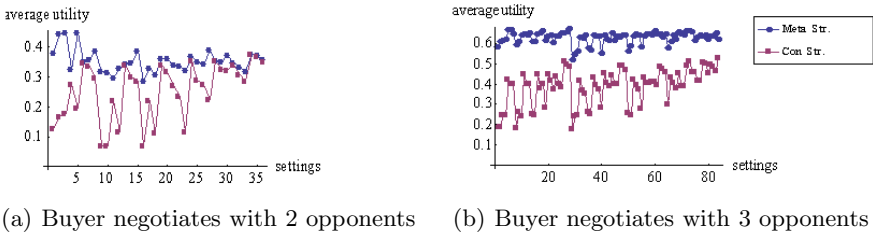
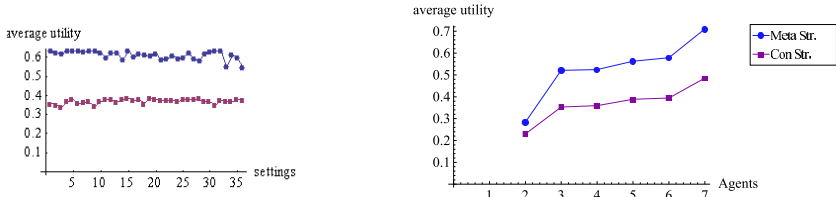


Fig. 2. Average utility subfigures (a) and (b)

Figure 2(a) shows the buyer’s average utility when the buyer agent negotiates with 2 seller agents (i.e., opponents). It shows that the average utility in all possible negotiation environments for the proposed meta-strategy is better than using the concession strategy, while Figure 2(b) shows the results when the buyer agent negotiates with 3 opponents ($\binom{9}{3} = 84$ different settings.) Again, the results shown in Figure 2(b) indicate that the meta-strategy performs better when compared to the concession strategy in case the buyer agent negotiates with 3 seller agents.

The results show that when the number of opponents increases, the difference in utility between the two strategies increases to due to the fact the more opponents would give more advantage to the buyer agent (compare between Figures 2(a), 2(b) and 3(a).) We also notice that the variation in the average utility for both strategies is smaller when the buyer agent negotiates with more opponents due to the fact that more opponents (compare between Figures 2(b) and 3(a)) means better competitive advantage and hence guarantees more stable results for the buyer agent.

Figure 3(b) shows the average utility versus the number of opponents. It is obvious that the number of opponents affects the utility regardless the negotiation strategy. At the same time, using the proposed meta-strategy outperforms the concession strategy.



(a) Buyer negotiates with 7 opponents (b) Avg. utility vs. number of opponents

Fig. 3. Average utility subfigures (a) and (b)

As a conclusion, the proposed meta-strategy outperforms the concession strategy regardless the negotiation environment and the number of opponents. As a final remark, the results do not show significant difference in the agreement rate between both strategies.

6 Conclusion and Future Work

This study proposes a meta-strategy as a coordination mechanism for the one-to-many negotiation form where a buyer agent negotiates with multiple seller agents over multiple continuous issues concurrently. After few rounds of negotiations, the buyer agent classifies the opponents into two groups according to their behaviors, a favorable group and unfavorable one. For the favorable group, the buyer agent uses the proposed meta-strategy, and for the unfavorable group, it uses the concession strategy to reduce the risk of not reaching an agreement. In addition, we propose the **IOG** as an offer generation method in which an agent concedes on the least ranked issues first in case of using concession strategy and concedes on the issues that probably will generate high utility values for the opponents first when the trade-off method is used. We test our proposed meta-strategy coordination mechanism under wide number of different negotiation environments. The results show that our meta-strategy is robust and produces better results than using the concession strategy in terms of the buyer agent's utility gain.

Future work will investigate the situation where agents negotiate over multiple objects/multiple issues and study possible techniques that can be used by the buyer agent negotiating with multiple sellers concurrently. In addition, the **IOG** method will be further investigated to test whether agents can reach near Pareto optimal solution when all agents use it as an offer generation mechanism.

References

1. An, B., Lesser, V., Irwin, D., Zink, M.: Automated Negotiation with Decommittment for Dynamic Resource Allocation in Cloud Computing. In: 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010), Toronto, pp. 981–988 (2010)

2. An, B., Sim, K.M., Miao, C.Y., Shen, Z.Q.: Decision making of negotiation agents using markov chains. *Multiagent and Grid Systems* 4, 5–23 (2008)
3. Cuihong, L., Giampapa, J., Sycara, K.: Bilateral negotiation decisions with uncertain dynamic outside options. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 36(1), 31–44 (2006)
4. Faratin, P.: Automated service negotiation between autonomous computational agents. PhD thesis, University of London (2000)
5. Faratin, P., Sierra, C., Jennings, N.R.: Using similarity criteria to make issue trade-offs in automated negotiations. *Artificial Intelligence* 142(2), 205–237 (2002)
6. Faratin, P., Sierra, C., Jennings, N.R.: Using similarity criteria to make issue trade-offs in automated negotiations. *Artificial Intelligence* 142, 205–237 (2002)
7. Fatima, S.: An agenda-based framework for multi-issue negotiation. *Artificial Intelligence* 152(1), 1–45 (2004)
8. Fatima, S.S., Wooldridge, M.J., Jennings, N.R.: Optimal Negotiation Strategies for Agents with Incomplete Information. In: Meyer, J.-J.C., Tambe, M. (eds.) *ATAL 2001. LNCS (LNAI)*, vol. 2333, pp. 53–68. Springer, Heidelberg (2002)
9. Jennings, N.R., Faratin, P., Lomuscio, A.R., Parsons, S., Wooldridge, M., Sierra, C.: Automated Negotiation: Prospects, Methods and Challenges. *Group Decision and Negotiation* 10, 199–215 (2001)
10. Mansour, K., Kowalczyk, R., Vo, B.Q.: Real-Time Coordination of Concurrent Multiple Bilateral Negotiations under Time Constraints. In: Li, J. (ed.) *AI 2010. LNCS (LNAI)*, vol. 6464, pp. 385–394. Springer, Heidelberg (2010)
11. Di Napoli, C.: Software Agent Negotiation for Service Composition. In: Nguyen, N.T., Grzech, A., Howlett, R.J., Jain, L.C. (eds.) *KES-AMSTA 2007. LNCS (LNAI)*, vol. 4496, pp. 456–465. Springer, Heidelberg (2007)
12. Nguyen, T., Jennings, N.: Managing commitments in multiple concurrent negotiations. *Electronic Commerce Research and Applications* 4(4), 362–376 (2005)
13. Nguyen, T.D., Jennings, N.R.: Concurrent bi-lateral negotiation in agent systems. In: *Proceedings of the Fourth DEXA Workshop on E-Negotiations* (2003)
14. Nguyen, T.D., Jennings, N.R.: Coordinating multiple concurrent negotiations. In: *The Third International Joint Conference on Autonomous Agents and Multi Agent Systems*, New York, USA, pp. 1062–1069 (2004)
15. Osborne, M., Rubinstein, A.: *A Course in Game Theory*. MIT Press, Cambridge (1994)
16. Rahwan, I., Kowalczyk, R., Pham, H.H.: Intelligent agents for automated one-to-many e-commerce negotiation. In: *Twenty-Fifth Australian Computer Science Conference*, Melbourne, Australia, pp. 197–204 (2002)
17. Ros, R., Sierra, C.: A Negotiation Meta Strategy Combining Trade-off and Concession Moves. *Autonomous Agents and Multi-Agent Systems* 12(2), 163–181 (2006)
18. Wong, T.N., Fang, F.: A multi-agent protocol for multilateral negotiations in supply chain management. *International Journal of Production Research* 48(1), 271–299 (2010)

Further Research of Generated Filters in Lattice Implication Algebra

Ling Guo¹, Yang Xu¹, and Shaokun Du²

¹ Intelligent Control Development Center, Southwest Jiaotong University,
Chengdu 610031, China
guo1987ling@163.com

² Basic Department, Hebei Engineering and Technical College,
Changzhou 061001, China

Abstract. In this paper we combining LHIA research the generated filters of lattice implication algebra. Firstly some new properties of lattice implication algebra are discussed. Then a structure of generated filter of two filters' union is obtained. Finally some elementary properties are given and a simple proof of $(F(L), \vee, \wedge)$ is a complete distributive lattice is also obtained.

Keywords: Lattice implication algebra, Filters, Generated filters.

1 Introduction

As is well known to us that certain information processing, especially inferences based on certain information, is based on the classical two-valued logic. Naturally, it is necessary to establish some rational logic systems as the logical foundation for uncertain information processing. For this reason, various kinds of non-classical logic systems have been extensively proposed and researched. In fact, non-classical logic has become a formal and useful tool for computer science to deal with uncertain information and fuzzy information. On the other hand, various logical algebras have been proposed as the semantic systems of non-classical logic systems, for example, Many-valued logic have always been a crucial direction in non-classical logic.

In order to research the many-valued logical system whose Propositional value is given in a lattice, Xu [1] (1993) established the concept of lattice implication algebras by combining lattices and implication algebras, and gave its elementary properties. The filter theory of the logical algebras plays an important role in studying these algebras and the completeness of the corresponding non-classical logics. At present, for the general development of algebra systems, the filter theory has been widely studied, and some important results have been obtained. Meanwhile, filter plays an important role in automated reasoning and approximated reasoning based on lattice implication algebra, Xu and Qin [2] introduced the notions of filter, implicative filter and generated filters in a lattice implication algebra, Liu [5] investigated more details about generated filters and got some properties. Xu and Qin [6] prove that $A(a)$ is the filter generated by a in a lattice H implication algebras, and got $A(a) \vee A(b)$ is a filter in a LHIA. And more properties of lattice implication algebras and lattice H implication algebras can be seen in [3,4,12,14,16]. In this paper, we study and discuss the

properties of the filters on lattice implication algebra, also other important characteristic of the filters are given. These works will become an important algebraic foundation for us to study the corresponding logical system and uncertain reasoning theories and methods. Furthermore, we focus on generated filters, research the generated filter of union of two filters. Using the new structure, we give a simple proof of $(F(L), \vee, \wedge)$ is a complete distributive lattice.

2 Preliminaries

Definition 2.1[1]. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a bounded lattice with the universal bounds O, I and an inverse involution $'$, $\rightarrow : L \times L \rightarrow L$ be a mapping. $(L, \vee, \wedge, ', \rightarrow, O, I)$ is called a lattice implication algebra (shortly as *LIA*), if the following conditions hold for any $x, y, z \in L$:

- $(I_1) x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$;
- $(I_2) x \rightarrow x = I$;
- $(I_3) x \rightarrow y = y' \rightarrow x'$;
- $(I_4) x \rightarrow y = y \rightarrow x = I$ implies $x = y$;
- $(I_5) (x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x$;
- $(L_1) (x \vee y) \rightarrow z = (x \rightarrow z) \wedge (y \rightarrow z)$;
- $(L_2) (x \wedge y) \rightarrow z = (x \rightarrow z) \vee (y \rightarrow z)$.

Theorem 2.1[3]. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, then for any $x, y \in L$, the following statements are equivalent:

- (1) For any $x, y, z \in L$, $x \rightarrow (y \rightarrow z) = (x \wedge y) \rightarrow z$;
- (2) For any $x, y, z \in L$, $(x \rightarrow (y \rightarrow z)) \rightarrow ((x \rightarrow y) \rightarrow (x \rightarrow z)) = I$;
- (3) For any $x, y \in L$, $x \rightarrow (x \rightarrow y) = x \rightarrow y$.

Definition 2.2. [4] A lattice implication algebras L is called a lattice *H* implication algebra (LHIA for short), if for any $x, y, z \in L$,

$$x \vee y \vee ((x \wedge y) \rightarrow z) = I .$$

Corollary 2.1. [4] Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, then the following statements are equivalent:

- (1) $(L, \vee, \wedge, ', \rightarrow, O, I)$ is a lattice *H* implication algebra;
- (2) For any $x, y \in L$, $x \rightarrow (x \rightarrow y) = x \rightarrow y$;
- (3) For any $x, y, z \in L$, $(x \rightarrow (y \rightarrow z)) = (x \rightarrow y) \rightarrow (x \rightarrow z)$;
- (4) For any $x, y, z \in L$, $x \rightarrow (y \rightarrow z) = (x \wedge y) \rightarrow z$;
- (5) For any $x, y \in L$, $(x \rightarrow (y \rightarrow z)) \rightarrow ((x \rightarrow y) \rightarrow (x \rightarrow z)) = I$

(6) For any $x, y \in L$, and $z \in [x, I]$, $z \rightarrow (x \rightarrow y) = x \rightarrow y$;

(7) For any $x, y \in L$, $(x \rightarrow y) \rightarrow x = x$.

Corollary 2.2. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice H implication algebra, then for any $x, y \in L$,

$$((x \rightarrow y) \rightarrow x) \rightarrow (x \rightarrow (x \rightarrow y)) = x \rightarrow y.$$

Proof. From corollary 2.1 we get

$$\begin{aligned} & ((x \rightarrow y) \rightarrow x) \rightarrow (x \rightarrow (x \rightarrow y)) \\ &= x \rightarrow (x \rightarrow y) \\ &= x \rightarrow y \end{aligned}$$

Corollary 2.3. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice H implication algebra, then for any $x, y \in L$,

$$(x \rightarrow y) \rightarrow (y \rightarrow x) = y \rightarrow x.$$

Proof. Since $(x \rightarrow y) \rightarrow (y \rightarrow x)$

$$\begin{aligned} &= y \rightarrow ((x \rightarrow y) \rightarrow x) \\ &= y \rightarrow x \end{aligned}$$

In a lattice implication algebra $(L, \vee, \wedge, ', \rightarrow, O, I)$, we define binary operation \otimes as follows: for any $x, y \in L$,

$$x \otimes y = (x \rightarrow y')'.$$

For $x, y \in L$, and $n \in N$, we denote that

$$a^0 = I, \quad a^1 = a, \quad a^{n+1} = (a^n) \otimes a.$$

And in what follows,

$$[a_1, a_2, \dots, a_n, x] \triangleq a_1 \rightarrow (a_2 \rightarrow (\dots \rightarrow (a_n \rightarrow x) \dots)).$$

Specially,

$$\begin{aligned} [a, x]^0 &\triangleq x, \\ [a, x]^1 &\triangleq [a, x] = a \rightarrow x, \\ [a, x]^n &\triangleq \underbrace{a \rightarrow (a \rightarrow (\dots \rightarrow (a \rightarrow x) \dots))}_n \quad (n \geq 2). \end{aligned}$$

Theorem 2.2. [3] In a lattice implication algebra $(L, \vee, \wedge, ', \rightarrow, O, I)$, the following hold:

- (1) $x \otimes y = y \otimes x$,
- (2) $(x \otimes y) \otimes z = x \otimes (y \otimes z)$,

- (3) $x \otimes y \leq x, x \otimes y \leq y,$
- (4) $O \otimes x = O, I \otimes x = x, x \otimes x' = O,$
- (5) $x \rightarrow (x \otimes y) = x' \vee y,$
- (6) $(x \rightarrow y) \otimes x = x \wedge y,$
- (7) If $x \leq a$ and $y \leq b,$ then $x \otimes y \leq a \otimes b,$
- (8) $x \otimes (a \vee b) = (x \otimes a) \vee (x \otimes b),$ and $x \otimes (a \wedge b) = (x \otimes a) \wedge (x \otimes b).$

3 Some Properties of Lattice Implication Algebras

Definition 3.1. [2] Let L be a lattice implication algebra, $J \subseteq L$ is said to be a filter of L , if it satisfies the following conditions:

- (1) $I \in J;$
- (2) For any $x, y \in L,$ if $x \in J$ and $x \rightarrow y \in J,$ then $y \in J.$

Corollary 3.1[2]. $J \subseteq L$ is said to be an filter of $L,$ if it satisfies the following conditions:

- (1) $I \in J;$
- (2) For any $x, y \in J$ and $z \in L,$ if $x \rightarrow (y \rightarrow z) \in J,$ then $z \in J.$

Corollary 3.2. $J \subseteq L$ is a filter of $(L, \vee, \wedge, ', \rightarrow, O, I)$ if and only if it satisfies the following conditions:

- (1) $I \in J;$
- (2) For any $a_i \in J, i = 1, \dots, n,$ if $[a_1, a_2, \dots, a_n, x] \in J,$ then $x \in J.$

Corollary 3.3. $J \subseteq L$ is a filter of $(L, \vee, \wedge, ', \rightarrow, O, I)$ if and only if it satisfies the following conditions:

- (1) $I \in J;$
- (2) For $a \in J, \exists n \in N^+,$ if $[a, x]^n \in J,$ then $x \in J.$

Proof. The two corollaries are easily proved by corollary3.1.

Theorem 3.1[2]. Let L be a lattice implication algebra, $J \subseteq L$ is a filter of L if and only if:

- (1) $I \in J;$
- (2) If $a \in J$ and $a \leq b,$ then $b \in J.$
- (3) If $a, b \in J,$ then $a \otimes b \in J.$

Theorem 3.2. Let L be a lattice implication algebra, $J \subseteq L$ is a filter of L if and only if:

- (1) $I \in J$;
- (2) If $a \in J$ and $a \leq b$, then $b \in J$.
- (3) If $a_i \in J, i = 1, \dots, n$, then $a_1 \otimes a_2 \otimes \dots \otimes a_n \in J$.

Proof. If J is a filter of L , then (1) and (2) hold. For (3) we can easily get by theorem 3.1.

Conversely, suppose that (1), (2) and (3) hold. If $x \in J$ and $x \rightarrow y \in J$, then we have $x \otimes (x \rightarrow y) \in J$. Hence we get $y \in J$ for $x \otimes (x \rightarrow y) = x \wedge y \leq y$, so J is a filter.

Corollary 3.4. Let L be a lattice implication algebra, $J \subseteq L$ is a filter of L if $a \in J$, $\forall n \in \mathbb{N}^+$, then $a^n \in J$.

Proof. This corollary is easily got by theorem 3.2.

Definition 3.2[2]. Let L be a lattice implication algebra, $J \subseteq L$ is said to be an implicative filter of L , if it satisfies the following conditions:

- (1) $I \in J$;
- (2) For any $x, y, z \in L$, if $x \rightarrow (y \rightarrow z) \in J$ and $x \rightarrow y \in J$, then $x \rightarrow z \in J$.

Theorem 3.3[2]. An implicative filter is a filter in a lattice implication algebra, but the converse is not true.

Definition 3.3[3]. Let L be a lattice implication algebra, $\alpha \in L$. Denote the interval $[\alpha, I]$

$$[\alpha, I] = \{x \mid x \in L, \alpha \leq x\}$$

of L as $A_{(\alpha)}$.

Definition 3.4. Let L be a lattice implication algebra and F a filter of L , $\forall a \in L$. We denote

$$F_a = \{x \in L \mid a \rightarrow x \in F\} .$$

Theorem 3.4. Let F be a filter of lattice implication algebra L , then F is an implicative filter if and only if for any $a \in L$, F_a is a filter of L .

Proof. Let F be an implicative filter, For any $a \in L$. Since $a \rightarrow I = I \in F$, so $I \in F_a$. If $x, x \rightarrow y \in F_a$, then $a \rightarrow x \in F$ and $a \rightarrow (x \rightarrow y) \in F$, Since F is an implicative filter, hence $a \rightarrow y \in F$, then $y \in F_a$. Therefore F_a is a filter of L .

Conversely, for any $a \in L$, F_a is a filter. $\forall x, y, z \in L$,

if $x \rightarrow (y \rightarrow z) \in F, x \rightarrow y \in F$. Then $y \rightarrow z \in F_a$ and $y \in F_a$. Since F_a is a filter, we have $z \in F_a$, i.e., $x \rightarrow z \in F$. Hence F is an implicative filter.

Definition 3.5[3]. Let L be a lattice implication algebra. A subset J of L is called a positive implicative filter of L if it satisfies:

- (1) $I \in J$;
- (2) For any $x, y, z \in L$, if $x \rightarrow ((y \rightarrow z) \rightarrow y) \in J$ and $x \in J$ implies $y \in J$.

Corollary 3.5[3]. In a lattice implication algebra, the concepts of implicative filters and positive implicative filters coincide.

Corollary 3.6. Let L be a lattice implication algebra, J_1, J_2 are filters of L , $J_1 \subseteq J_2$. If J_1 is a positive implicative filter, so is J_2 .

Theorem 3.5[3]. Let L be a lattice implication algebra and J a filter of L . Then J is a positive implicative filter if and only if $(x \rightarrow y) \rightarrow x \in J$ implies $x \in J$ for any $x, y \in L$.

Theorem 3.6. In a lattice implication algebra L , the following conditions are equivalent:

- (a) $\{I\}$ is a positive implicative filter,
- (b) Every filter of L is a positive implicative filter,
- (c) $A_{(a)} = \{x \mid x \in L, a \leq x\}$ is a positive implicative filter,
- (d) $(x \rightarrow y) \rightarrow x = x$ for al $x, y \in L$,

Proof. (a \Rightarrow b): By corollary 3.6 is clear.

(b \Rightarrow c): Since $\{I\}$ is a positive implicative filter, then $\{I\}$ is an implicative filter. Since $a \leq I$, we have $I \in A_{(a)}$. Let $x, x \rightarrow y \in A_{(a)}$, then $a \rightarrow x = I, a \rightarrow (x \rightarrow y) = I$.

Then we get $a \rightarrow y = I$ and so $a \leq y$. Therefore $y \in A_{(a)}$. Hence $A_{(a)}$ is a filter and by (b), $A_{(a)}$ is a positive implicative filter.

(c \Rightarrow d): We have $(x \rightarrow y) \rightarrow x \in A_{((x \rightarrow y) \rightarrow x)}$, By (c) $A_{((x \rightarrow y) \rightarrow x)}$ is a positive implicative filter, then by theorem 3.5 $x \in A_{((x \rightarrow y) \rightarrow x)}$ and so $(x \rightarrow y) \rightarrow x \leq x$. We have $(x \rightarrow y) \rightarrow x \geq x$. thus $(x \rightarrow y) \rightarrow x = x$.

(d \Rightarrow a): The proof is easy by theorem 3.5.

4 Further Research on Generated Filters

In reference [4], Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, it is easy to verify that the intersection of filters of L is also a filter, .Suppose $A \subseteq J$, the least filter containing A is called the filter generated by A and denoted by $[A]$. It is trivial to verify that

$$[A] = \bigcap \{B \mid A \subseteq B \subseteq L, B \text{ is a filter}\} .$$

In what follows, $[\{a\}]$ is denoted by $[a]$.

Example 4.1. As for the lattice implication algebra defined in example2.1.4 [3], it is routine to check that if $A = \{b, c\}$, then $[A] = \{b, c, I\}$, and $[a] = \{a, I\}$ is the filter generated by $\{a\}$.

Theorem 4.1. L is a Lukasiewicz implication algebra. Then

$$[A] = \begin{cases} L & A \neq I; \\ I & A = I. \end{cases}$$

Proof. As we know that all Lukasiewicz implication algebras have only trivial filters. so it's obvious that the theorem is true.

Theorem 4.2 [2]. Let L be a lattice implication algebra, $\emptyset \neq A \subseteq L$. Then

$$[A] = \{x \mid x \in L, \text{there exist } a_1, \dots, a_n \in A, n \in \mathbb{N}^+, \text{s.t.}, [a_1 \cdots a_n, x] = I\}.$$

Corollary 4.1. Let L be a lattice implication algebra, $\emptyset \neq A \subseteq L$. Then

$$[A] = \{x \mid x \in L, \text{there exist } a \in A, n \in \mathbb{N}^+, \text{s.t.}, a^n \rightarrow x = I\}$$

Theorem 4.3[3]. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, J is an implicative filter of L , $a \in L$, then $[J \cup a] = \{x \in L \mid a \rightarrow x \in J\}$.

Proposition 4.1 [2]. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra. Then L is a lattice H implication algebra if and only if every filter of L is also an implicative filter.

Corollary 4.2. If J is a filter of lattice H implication algebra L , $a \in L$, then

$$[J \cup a] = \{x \in L \mid a \rightarrow x \in J\}.$$

In a lattice H implication algebra L , we define a sub-structure of L , $\forall a, b \in L$, $A(a, b) = \{x \in L \mid [a, b, x] = I\}$. And we can easily get the concepts of $A(a, I)$ and $A(a)$ are coincide.

Theorem 4.4. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice H implication algebra, then $A(a, b)$ is a filter of L .

Proof. Firstly, $I \in A(a, b)$, in fact $[a, b, I] = a \rightarrow (b \rightarrow x) = (a \wedge b) \rightarrow I = I$, suppose $x, x \rightarrow y \in A(a, b)$, it follows that $a \rightarrow (b \rightarrow x) = I$ and $a \rightarrow (b \rightarrow (x \rightarrow y)) = I$, since $a \rightarrow (b \rightarrow (x \rightarrow y)) = a \rightarrow ((b \rightarrow x) \rightarrow (b \rightarrow y)) = (a \rightarrow (b \rightarrow x)) \rightarrow (a \rightarrow (b \rightarrow y)) = I$, then $a \rightarrow (b \rightarrow y) = I$, that is $y \in A(a, b)$. Thus $A(a, b)$ is a filter.

Theorem 4.5. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice H implication algebra, $a, b \in L$, then $[a \wedge b] = A(a, b)$.

Proof. Firstly, from theorem 4.4 we know that $A(a,b)$ is a filter of L . Now we prove $A(a,b)$ is the least filter containing $a \wedge b$, it's obviously that $a \wedge b \in A(a,b)$, let $a \wedge b \in B$, B is a filter. $\forall x \in A(a,b)$, then $a \rightarrow (b \rightarrow x) = (a \wedge b) \rightarrow x = I$, so $(a \wedge b) \rightarrow x \in B$, then $x \in B$, so $A(a,b) \subseteq B$, i.e., $[a \wedge b] = A(a,b)$.

Theorem 4.6[3]. Let L be a lattice implication algebra, $\emptyset \neq A \subseteq L$. Then

$$[A] = \{x \mid x \in L, \text{there exist } a_1, \dots, a_n \in A, n \in N^+, \text{s.t.}, [a_1 \dots, a_n, x] = I\}.$$

Corollary 4.3. Let L be a lattice implication algebra, For any $a \in L$. Then

$$[a] = \{x \mid x \in L, \text{there exist } n \in N^+, \text{s.t.}, a^n \leq x\}.$$

In the following, we let $F(L)$ be the set of all filters of L .

Theorem 4.7. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, $A, B \in F(L)$, then the following hold:

- (1) if $a \in L, [a] \subseteq A$ if and only if $a \in A$,
- (2) $[A \cup B] = \{x \in L \mid \exists a \in A \text{ and } b \in B, \text{ such that } a \otimes b \leq x\}$.

Proof. Firstly, we prove the necessary of (1), since $[a] \subseteq A$, if $a \notin A$, then $[a] \not\subseteq A$, that is a contradiction. Then the sufficiency, since $a \in A$, and A is a filter, so $[A] = A$, then we get $[a] \subseteq A$.

(2) Suppose $C = \{x \in L \mid \exists a \in A \text{ and } b \in B, \text{ such that } a \otimes b \leq x\}$, Firstly, we prove C is a filter, in fact, by $I \in A, I \in B$, then $I \in C$, suppose $x, x \rightarrow y \in C$, then there are $a_1, a_2 \in A, b_1, b_2 \in B$, such that $a_1 \otimes b_1 \leq x, a_2 \otimes b_2 \leq x \rightarrow y$, by the properties of \otimes , we get $(a_1 \otimes b_1) \otimes (a_2 \otimes b_2) = (a_1 \otimes a_2) \otimes (b_1 \otimes b_2) \leq x \otimes (x \rightarrow y) \leq y$. For $a_1 \otimes a_2 \in A, b_1 \otimes b_2 \in B$, that is $y \in C$, so C is a filter.

Secondly, we prove that C is the least filter containing $A \cup B$, if $x \in A \cup B$, then $x \in C$. In fact,

- (a) If $x \in A \setminus B$, then $\exists I \in B$, such that $x \otimes I \leq x$, then $x \in C$,
- (b) If $x \in B \setminus A$, this is similar with (a),
- (c) If $x \in A$ and $x \in B$, then $x \otimes x \leq x$, that is $x \in C$.

Further $A \cup B \subseteq C$. if $A \cup B \subseteq D$, and D is a filter, $\forall x \in C$, then $\exists a \in A \subseteq D, b \in B \subseteq D$, such that $a \otimes b \leq x$, then $x \in D$, then $C \subseteq D$, so the conclusion is hold.

Corollary 4.4. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, J is a filter of $L, a \in L$, then

$$[J \cup a] = \{x \in L \mid \exists j \in J \text{ and } n \in N, \text{ such that } j \otimes a^n \leq x\}.$$

For Theorem 3.2.10 in [4], the proof is too complex to understand. We give the simpler proof for this theorem as follows.

Theorem 4.8. Let $(L, \vee, \wedge, ', \rightarrow, O, I)$ be a lattice implication algebra, then $(F(L), \vee, \wedge)$ is a complete distributive lattice, and for any $A, B \in F(L)$, $A \wedge B = A \cap B$, $A \vee B = [A \cup B]$.

Proof. In the following, we only need to prove that the distributive law holds, *i.e.* $A \cap [B \cup C] = [(A \cap B) \cup (A \cap C)]$ hold for any $A, B, C \in F(L)$. In fact, it's obviously that $A \cap [B \cup C] \supseteq [(A \cap B) \cup (A \cap C)]$ followed. Conversely suppose $x \in A \cap [B \cup C]$, then $x \in A$ and $x \in [B \cup C]$, then there exists $b \in B, c \in C$, such that $b \otimes c \leq x$, since $A \cap B$ and $A \cap C$ are filters, so $x \vee b \in A \cap B$, $x \vee c \in A \cap C$, and $(b \vee x) \otimes (c \vee x) = (b \otimes c) \vee (x \vee c) \vee ((b \vee x) \otimes x) \leq x$, then $x \in [(A \cap B) \cup (A \cap C)]$, further $A \cap [B \cup C] = [(A \cap B) \cup (A \cap C)]$. Thus $(F(L), \vee, \wedge)$ is a complete distributive lattice.

5 Conclusion

In this paper, we study and discuss the properties of the filters of lattice implication algebra, also the characteristic of the filters are given. And the generated filters are studied by using the properties of the filters, and some elementary properties of generated filters are given, research the generated filter of union of two filters, and give the structure of it. Also we use this structure prove $(F(L), \vee, \wedge)$ is a complete distributive lattice, which is simpler method. The further research will be concentrated on the generated filters on product lattice implication algebra.

Acknowledgments. This work is partially supported by the National Natural Science Foundation of China (Grant No. 60875034).

References

1. Xu, Y.: Lattice implication algebras. J. Southwest Jiaotong University 28, 20–27 (1993) (in Chinese)
2. Xu, Y., Qin, K.Y.: On filters of lattice implication algebras. J. Fuzzy Math. 1, 251–260 (1993)
3. Xu, Y., Ruan, D., Qin, K.Y.: Lattice-Valued Logic. Springer, Heidelberg (2003)
4. Xu, Y., Qin, K.Y.: Lattice H implication algebras and lattice implication algebra classes. J. Hebei Mining and Civil Engineering Institute 3, 139–143 (1992)
5. Liu, J., Xu, Y.: Filters and structure of lattice implication algebras. Chinese Science Bulletin 42(18), 1517–1520 (1997)
6. Xu, Y., Qin, K.Y., Roh, E.H.: On prime filters of lattice H implication algebras. Int. J. Fuzzy Math. 9(3), 583–588 (2001)
7. Haveshki, M., Saeid, A.B., Eslami, E.: Some types of filters in BL-algebras. Soft. Comput. 10, 657–664 (2006)

8. Hu, M.D., Lou, Z.G.: Normal MP-filters and prime filters of an implication lattice. *J. Shandong University (Natural Science)* 43(10), 31–35 (2008) (in Chinese)
9. Li, J.C., Zhang, W.X.: Quasi-Fuzzy Valuations on HFI-Algebras. *Fuzzy Systems and Mathematics* 12(4), 1–3 (2000) (in Chinese)
10. Huang, T.M., Xu, Y., Zhao, H.L., Wu, J.L.: Lattice Order Introduction and the Application. *Journal of Southwest Jiaotong University* (1998) (in Chinese)
11. Li, H.M., Pei, Z., Meng, D., Xu, Y.: Some Remarks on Structure of Lattice Implication Algebras. *Fuzzy Systems and Mathematics* 17(4), 20–24 (2003) (in Chinese)
12. Jun, Y.B., Xu, Y., Qin, K.Y.: Positive implicative and associative filters of lattice implication algebras. *Bull. Korean Math. Soc.* 35(1), 53–61 (1998)
13. Xu, Y., Qin, K.Y.: Fuzzy lattice implication algebras. *Southwest Jiaotong University* 2, 121–127 (1995)
14. Zhu, Y.Q., Xu, Y.: On filter theory of residuated lattices. *Information Sciences* 180, 3614–3632 (2010)
15. Dilworth, R.P., Ward, M.: Residuated lattices. *Trans. Am. Math. Soc.* 45, 335–354 (1939)
16. Xu, Y., Qin, K.Y.: The(S)condition in lattice implication algebras. *J. Decision and Operation* 2, 2039–2044 (1992)
17. Qin, K.Y.: Study of lattice-valued logic system and its application based on lattice implication algebra, Ph.D. thesis. Southwest Jiaotong University, Chengdu, China (1996)
18. Turunen, E.: Boolean deductive systems of BL-algebras. *Arch. Math. Logic* 40, 467–473 (2001)

Dynamic Task Allocation and Action Coordination under Uncertain Environment

Chengli Liu, Wei Zeng^{*}, Hongtao Zhou, Lei Cao, and Yang Yang

Department of Control Science and Engineering,
Huazhong University of Science & Technology, Wuhan 430074, China
zengwei@mail.hust.edu.cn

Abstract. Under complex, dynamic and uncertain environment, tasks in multi-agent system need to be distributed to agents dynamically and agents need to cooperate to complete tasks assigned dynamically. This paper proposes a dynamic task allocation model based on game theory and dynamic coordination in task execution based on coordination graph at each time step. The synthesized model is solved by reinforcement learning. The detailed algorithm is illustrated with an example and experimental results show that the method is an effective solution for dynamic task allocation and execution coordination under uncertain environment.

Keywords: task allocation, coordination graph, Q-learning, game theory.

1 Introduction

Under complex, dynamic and uncertain environment, task planning and execution are critic research issues in the domain of emergency rescue, military mission planning and multi-robot task planning. At present, task allocation approaches mainly focus on linear programming [1][2], market mechanism, and et al, while dynamic coordination methods in task execution include joint action learning [3], shared Q-value table, contract net and et al. However, in linear programming method, scalability of the number of agents in multi-agent system is poor due to matrix computation. In market mechanisms method and contract net protocol, multi-agent system requires large amounts of real-time communication among agents. In joint action learning, each agent needs to store an individual Q-function based on its own and others' actions, which is likely to cause the curse of dimensionality of state space.

In practice we need to simultaneously consider task allocation and dynamic task execution coordination. Therefore, we propose dynamic task allocation based on game theory as the dynamic nature of the environment. In the task execution, we apply coordination graph to represent the potential influences between agents and variable elimination method on coordination graph to compute optimal joint actions at specific states. At every decision-making period, the agents get optimal task assignment and optimal action strategies by reinforcement learning algorithm. The proposed method improves convergence efficiently and is illustrated with a predator-prey example.

^{*} Corresponding author.

2 Modeling

2.1 Task Allocation

In multi-agent system, there are n tasks which need to be assigned to the m agents. A rational agent selects the optimal task so as to maximize its own utility, according to forecasts of the task selection of the other agents. This section proposes the dynamic task allocation model on the basis of game theory. At each decision-making moment, there may form a game situation according to circumstances of the agent and decisions of the others. In order to solve the Nash equilibrium point, each agent under game situation selects an appropriate task consistent with its maximal expected utility. Without knowing the actual selections of the others, the agent concerns about its own expected return. Nash equilibrium of mixed strategy is based on the largest expected return of each participant.

$$\begin{aligned}
 &\text{In game } G = \{ \Gamma; S_1, \dots, S_n; \mu_1, \mu_2, \dots, \mu_n \} \\
 &V_i (S_i^*, S_{-i}^*) \geq V_i (S_i, S_{-i}^*) \quad \forall S_i \in \sum_i \quad i
 \end{aligned} \tag{1}$$

If the equation (1) holds for all agent $i = 1, 2, \dots, n$. The mixed strategy combination $S_1^*, S_2^*, \dots, S_n^*$ is a Nash equilibrium of mixed strategy. \sum_i is the total strategy space. Γ is the set of involved individuals. S_i is strategy space of the individual i . V_i is expected return of agent i . S_i^* and S_{-i}^* respectively optimal strategy of the individual i and optimal strategy of all individuals except individual i .

2.2 Coordination in Task Execution

All agents in multi-agent system may potentially interact with each other in some states. For example, two agents move into the same grid at the same time, which may result in conflicts. Or an agent may not complete task alone without the other agent’s help, due to limited ability. A coordination graph (CG) [4] represents direct or indirect dependency between agents. A node represents an agent, while an edge in the graph defines a possible dependency or interaction between two agents at some state. The agents interconnected need to coordinate their actions. Figure 1 shows a possible CG for a 4-agent problem. Rules [5] in the form $\langle \rho; c: v \rangle$ define v as value which is added to global payoff when the pair of state and action happens, where c (the context) is defined as an element from the set of all possible combinations of the state and action.

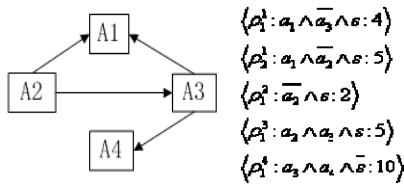


Fig. 1. Coordination graph with rules

Variable elimination [6] is a way to select optimal joint actions at any particular state. Eliminate nodes (or agents) in coordination graph one by one until only one agent is left. Then the final agent selects its optimal action corresponding to maximal v , the elimination is reversed. The other agents gradually get their own optimal actions one by one based on their own adjacent agents' action selection at context-specific states.

In the model, task selection and coordination are respectively solved by Q-learning and rule-learning. The rule learning is shown as follows:

$$\rho_k^j(s_i, a_i) \leftarrow \rho_k^j(s_i, a_i) + \alpha \sum_{l \in \{i, j\}} \frac{R_l(s, a) + \gamma Q_l(s'_i, s_t, t, a_i^*) - Q_l(s_i, s_t, t, a_i)}{|\Gamma(l)|} \quad (2)$$

where $\rho_k^j(s_i, a_i)$ denotes rule j of agent k at the state-action combination (s_i, a_i) . $\Gamma(l)$ denotes agents involved in rule except agent l .

Q-value of each agent is updated by Q-learning algorithm.

2.3 Algorithm

Based on the model above, the algorithm flowchart is shown as follows:

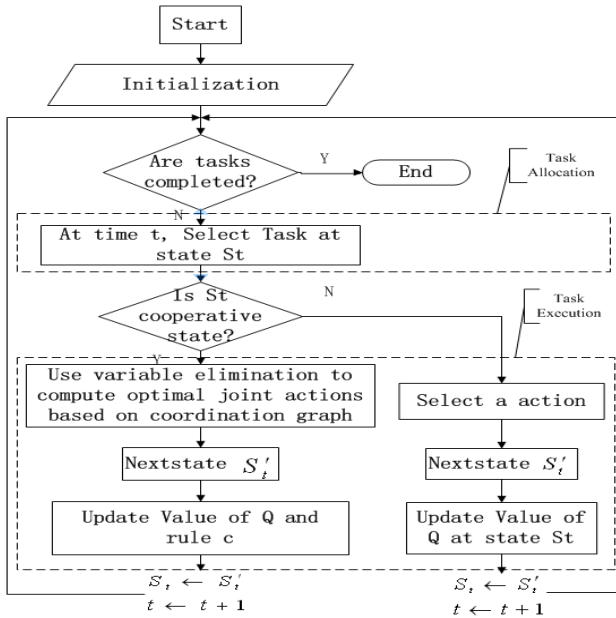


Fig. 2. Algorithm flowchart

Initialization: Initialize states of agent 1, 2 ... m and target 1, 2, ..., n , $Q_i(s_{i_s}, s_{i_t}, t, a)$ of agent $i, i=1, 2, \dots, m$, where s_{i_s}, s_{i_t} respectively are state of agent i and state of target t .

Task Allocation: Based on game theory, agent i selects task consistent with the maximal expected Q -value.

Task Execution: The agent takes an action from optimal joint actions at cooperative state, while takes some action at non-cooperative state. Update Q -value and rule value corresponding to state respectively based on formula (2) and (3).

3 Experimental Results

The proposed algorithm is applied to a cooperative multi-agent predator-prey problem. There are four agents (predators) and two tasks (preys) in the example. The tasks need be allocated and agents coordinate their actions in real-time manner. Agents in this experiment are supposed to move in a 10x10 toroidal grid world in Fig.3, and may move in four directions or stay in place. The prey is captured when one predator enters its square, while another predator is adjacent to prey, as shown in Fig.4.

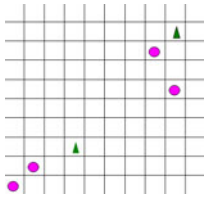


Fig. 3. Predator-Prey domain

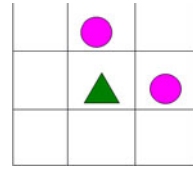


Fig. 4. Possible capture position for two predators

Each predator i gets a reward $R_i = 37.5$ when it succeeds to capture the prey with the other agent, a reward $R_i = -25.0$ when it collides with another predator, and a reward $R_i = -10.0$. When predator i moves into grid of the prey without the other predator's support, the reward is -0.5 in all other cases.

Predators should coordinate when either of two conditions holds:

1. The Manhattan distance between them is less than or equal to two cells.
2. Both predators are assigned to the same prey.

We define the rules at cooperative state as following:

$$\langle \rho_2^1; agent_1(0,1) \wedge agent_2(1,0) \wedge a_1 = none \wedge a_2 = west : 75 \rangle \quad (3)$$

Relative location of agent 1 is (0, 1) if condition 1 holds. If condition 2 holds, (0, 1) is local state of agent 1.

We conduct two experiments. One is the synthesized model mentioned above, and the other is fixed task assignment. We set the learning rate $\alpha = 0.1$, discount rate $\gamma = 0.9$. The rule value of cooperative state is initialized by 75, non-cooperative state is 37.5 and Q -value is 37.5. Average reward is measured after 10000 learning episodes.

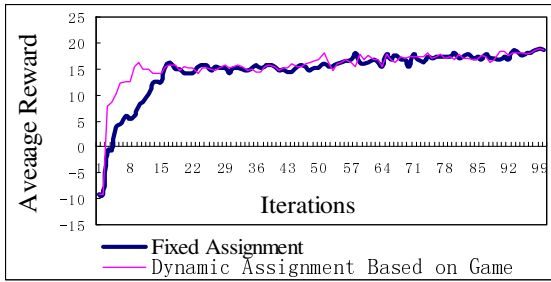


Fig. 5. Comparison of dynamic assignment based on game and fixed assignment

As may be seen from these results in Figure 5, dynamic task allocation performs better compared to fixed assignment. In task allocation, the agents can dynamically select task consistent with maximal Q-value at each time step so that expected return of the agent can converge faster than fixed task assignment after 10000 learning episodes. Experimental results show the synthesized model mentioned above is effective under dynamic and uncertain environment.

4 Conclusions

In this paper we propose a synthesized model of dynamic task assignment and action coordination in task execution. This method can be applied widely in real-time strategy games, disaster response, fire fighting, and et al, where multiple agents and tasks are involved. Experimental result shows that the synthesized model improves convergence efficiently. There are several research directions in the future works. A possible extension to the task allocation is to adopt cooperative game theory.

Acknowledgements. This research is supported by the Natural Science Foundation of China (Grant No. 70971048).

References

1. Gerkey, B.P., Matar, M.J.: A framework for studying multi-robot task allocation. In: Proceedings of the Multi-Robot Systems, Washington, USA, pp. 15–26 (2003)
2. Gerkey, B.P., Mataric, M.J.: A formal analysis and taxonomy of task allocation in multi-robot systems. *International Journal of Robotics Research* 23(9), 939–954 (2004)
3. Claus, C., Boutilier, C.: The dynamics of reinforcement learning in cooperative multiagent systems. In: Proc. 15th Nation. Conf. on Artificial Intelligence, Madison, WI
4. Jelle, R.K., Nikos, V.: Collaborative Multiagent Reinforcement Learning by Payoff Propagation. *Journal of Machine Learning Research* 7, 1789–1828 (2006)
5. Jelle, R.K., Nikos, V.: Sparse Cooperative Q-learning. In: Proceedings of the 21st International Conference on Machine Learning, Ban, Canada (2004)
6. Reinoud, E.: Anytime algorithms for multi-agent decision making, June 4 (2004)

Representation and Acquisition of Feature Value of Learner Model in Adaptive Learning System

Bing Jia, Yongjian Yang, and Jun Zhang

School of Computer Science, Jilin University, 2699, Qianjin Street,
130012, Changchun, China

Bing_jia@qq.com, yyj@jlu.edu.cn, june.zhang@163.com

Abstract. This paper was carried out mainly for the problem of learner model in adaptive learning system, such as the unscientific attention dimension, poor calculated representation method and single and subjective method to obtain feature values. We have put forward a new learner model to achieve self-organization of learning resources on the basis of learning goals and learner's personal conditions. It includes three feature items such as knowledge level, cognitive ability and preferences. We respectively introduce the representation and acquisition of feature value of learner model in detail. After that, we propose a push mechanism of learning resources. Experimental results show this learner model is effective and practical in the application.

Keywords: adaptive learning system, advanced learner model, presentation, acquisition, feature value.

1 Introduction

Learner Model(LM for short) is the representation of the learners' cognitive states [1]. It reflects the learners' learning progress, knowledge proficiency, misconceptions and the gap between the desired goals[2]. In adaptive learning system, LM is a single input system[3]. Currently there are three representative researches on it. We analysis and summarize these models from different prospects which are focusing on dimension, representation and acquiring method[4], Beginning with the prospective of focusing on dimension, it only concerns about the learners' interests, psychological features etc, however ignores such features as the learners' original knowledge level and cognitive ability which are relevant to the learning resources' recommending. Then, from the prospective of representation method, it generally uses the qualitative description method which is inconvenient in calculating and reasoning. In the end, from the prospective of acquiring method, it all uses single subjective methods such as questionnaire and scale, which made the acquisitive features inaccurate. In this article, in order to provide effective support to adaptive learning, we mainly put our focus on the research of the three above problems, aiming at researching the basic components and the relevant representation and acquiring method of the LM.

2 The Related Theoretical Basis

LM refers to produce one reliable way of expression to demonstrate what the student understood and could do, what he does not understand and could not do, what he wants to do as well as he should do[5].The basic idea of Adaptive Learning System is a student-centered, and a targeted study guide accord to the student mastery and acceptance of knowledge. It has the following features: adaptability, autonomy and resources construction.

LM is an abstract description and representation of learner's features information[6]. So how to determine the features of LM? First, Gagne[7]believes that any kind of learning new knowledge and skills are based on the knowledge already learned and subordinated to. Second, learner's cognitive capacity reflects his study ability (Speed, way to accept knowledge, etc). In this paper , we construct a advanced learner model (ALM for short)from three features of learner, such as knowledge level, cognitive ability and preferences.

3 Representation of ALM

3.1 Knowledge Level

Knowledge Level refers to a set of learner's existing knowledge under a certain goal. It has two parameters, one is knowledge point(KP for short), the other is mastering. We use k_i to represent the i_{th} KP, And we use k_i to represent learner's mastering of the i_{th} KP, h_i represents learner's mastering of the i_{th} KP. For the mastering, when the learner have no idea of the KP, we use "0" to represent the situation. For others, we divide the mastering into six levels based on Bloom's taxonomy of cognitive objectives, such as knowledge, comprehension, application, analysis, synthesis, evaluation, mapping with 1~6.

To sum up, we use "Knowledge-how" to represent the set.

$$\text{Knowledge-how} = \{(k_1, h_1), \dots, (k_i, h_i), \dots, (k_n, h_n)\} \quad (1)$$

where k_i represents the i_{th} KP, n represents the total of KPs, h_i represents learner's mastering of the i_{th} KP, $h_i \in H$, $H = \{0, 1, 2, 3, 4, 5, 6\}$, where "0" represents the situation that the learner have no idea of the KP, 1~6 respectively represents educational objectives proposed by Bloom[9](knowledge, comprehension, application, analysis, synthesis , evaluation)

3.2 Cognitive Ability

Cognitive Ability(CA for short) is individual's ability possessed while reconstructing and employing knowledge. We can describe the level of learners' ability by depicting the cognitive status of learners[10]. It also has two parameters, one is the type of CA, The other is the level value of CA. The type of CA should be divided into eight capability based on Gardner's multiple intelligences theory of cognitive abilities such as

Inductive capacity, memory, observation, abstraction ability, analytical ability, calculation ability, imagination, and logical reasoning ability. And in turn be expressed with a_1 to a_8 . The level value can be represented by l_i . As the ability to change is continuous, l_i is defined as the value greater than 0 and less than 1.

We use Ability-level to represent the set of learner's CA, so,

$$\text{Ability-level} = \{(a_1, l_1), \dots, (a_i, l_i), \dots, (a_n, l_n)\} \quad (2)$$

where a_i is the i_{th} CA, l_i is the level of the i_{th} CA, n is the total of CAs.

3.3 Preferences

Preferences are defined as demonstrated interest, hobby and other information in learning. Through the analysis, we preferred preferences information into the background material preferences, learning strategy preferences, learning time preferences, system features preferences and presentation preferences. In turn is expressed as P_b , P_s , P_t , P_f , P_p . They are five concept sets, A concept is a pair of (c, σ) , where c is a concept, and σ is a coefficient denotes the preference degree of the concept. For the value, take the presentation of resources for example, it can be taken the value of a set of $\langle \text{text}, \sigma_1 \rangle \langle \text{flash}, \sigma_2 \rangle$.

4 Acquisition of Feature Value of ALM

4.1 Estimation Method of Knowledge Level

In order to obtain a collection of learner's knowledge level "Knowledge-how", we need to take two steps, first is to get the set of the knowledge points (KPS for short) that learners have mastered, and the second is to obtain the mastery of KPS.

(1) Access to the set of KPS

For the set of KPS, we give the acquiring flow chart as shown in Fig. 1. The basic idea is to find its precursor according to the set of goal knowledge points, and estimate whether the value of the learners' master degree of these precursors is "puzzle", if not, end the process and record this KP. if so, carry on finding this knowledge point's precursor.

(2) Estimate the degree of learner's mastering of a KP

There are two effective methods of reasoning such as Bayes theorem and fuzzy logic. This paper mainly introduces the application of fuzzy logic in estimating the degree of learner's mastering of a certain KP [11]. According to Formula(1), we use these seven ranks (1~6) to express the degree of learner's mastering of a certain KP. The membership of each rank is expressed as $\mu_k(i), i=1\sim 6$. Express the degree of learner's mastering of a certain KP by fuzzy set K , so $K = \{i \mid (i, \mu_k(i)), i \in H\}$, where, $0 \leq \mu_k(i) \leq 1, \sum \mu_k(i) \leq 1$. To a KP, the highest level of membership is the level that the learner have achieved currently about this KP.

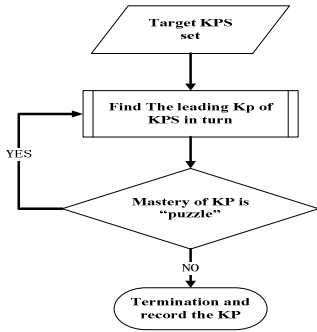


Fig. 1. Access to the set of KPS

Table 1. Changing Rules of Membership

If	Rule	Type
$i < m$	$\mu_k(i) = \mu_k(i) - \mu_k(i)q + \mu_k(i-1)q$	up rule
$i = m$	$\mu_k(i) = \mu_k(i) + \mu_k(i-1)q + \mu_k(i+1)q$	up rule
$i > m$	$\mu_k(i) = \mu_k(i) - \mu_k(i)q + \mu_k(i+1)q$	down rule

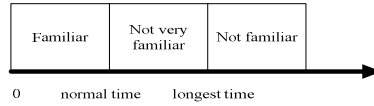


Fig. 2. Familiar degrees with the questions

We define the changing rules of membership of fuzzy set “K” as Table1, assume “m” is the rank of test questions, $i=1\sim 6$.

The basic principle of changing rule: when learner do a certain rank of test question correctly, adopt up rule to the rank and the following, and adopt down rule to the above. When learner do test question wrongly, the membership does not change. using “ Δ ” to represent the variation of $\mu_k(i)$, $\Delta=(\mu_k(i+1)- \mu_k(i))q$, so parameter “q” decides the scope of membership changing , It decides the declining speed of learner’s mastering of KP. The determination of parameter “q” should consider the following two aspects: the difficulty of the question and the degree of learner’s familiarity with the question. The difficulty of the question is expressed as “d”, where $0 < d < 1$.the degree of learner’s familiarity with the question will be obtained by the speed of the learner to answer the question correctly. We stipulate two time sections: normal time and longest time. If the learner answer the question correctly in normal time range, indicate that he is familiar with the knowledge. If he answers the question correctly between the normal time and the longest time, indicate that he is not very familiar with it. If beyond the longest time, indicate that he is not familiar with it. As Fig. 2 shows, Then, use the function to express the familiar degrees with the questions ,we can obtain the Definition 1.

Definition 1: Assume that α is the normal time, γ is the longest time, $\beta=(\alpha+\gamma)/2$.The function of the degree of learner’s familiarity with the question $F(t)$ is defined as: (where α, γ of each question can be different).

$$F(t)=1-S(t,\alpha,\beta,\gamma)=\begin{cases} 1 & \text{if } t \leq \alpha \\ 1-2\left(\frac{t-\alpha}{\gamma-\alpha}\right)^2 & \text{if } \alpha \leq t \leq \beta \\ 2\left(\frac{t-\gamma}{\gamma-\alpha}\right)^2 & \text{if } \beta \leq t \leq \gamma \\ 0 & \text{if } \beta \leq t \leq \gamma \end{cases}$$

Definition 1. Familiar degrees with the questions

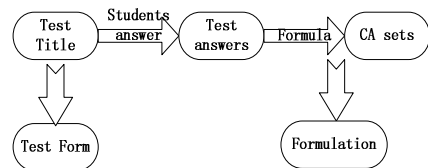


Fig. 3. The steps to obtain CA

So that $q = c^d * F(t)$, c is a constant larger than 1, $0 < d < 1$. When the learner does the question correctly in the normal time range, $F(t)=1$, q become large. When the learner does the question correctly exceed the normal range, $F(t)$ becomes small, q become small too. It proves that the learner's mastering degree of KP is changing slower When exceeds the longest time, $F(t)=0$, $q=0$. Similarly, q increases with increasing d . When the learner does a difficulty bigger question correctly, It proves he can learn this content well, also proves his mastering degree of KP is rising faster.

4.2 Estimation Method of Cognitive Ability

We can obtain the learner's CA by testing. The steps as shows in Fig. 3. We design a form of every test is : $\text{Test}(i) = (A(i), T(i), \eta, \lambda, \Phi: Q)$, A series of test questions constituted a test paper. $\text{TEST} = \{t_1 = (A_1, T_1, \eta_1, \lambda_1, \Phi_1 : Q_1), \dots, t_n = (A_n, T_n, \eta_n, \lambda_n, \Phi_n : Q_n)\}$. Then obtain the answers of the test questions, $\text{ANSWER} = \{\text{answer}(t_1, \phi'_1), \dots, \text{answer}(t_i, \phi'_i), \dots, \text{answer}(t_n, \phi'_n)\}$. So, we can define the level of the i_{th} CA as follow:

$$\text{Ability-how } (A_i) = \frac{\sum_{j=1}^n \left[\left(\frac{\phi'_j \cap \phi_j}{\phi_j} \right) \times \eta_j \times \lambda_j \right]}{\sum_{j=1}^n (\eta_j \times \lambda_j)} \quad (3)$$

where $(\phi'_j \cap \phi_j) / \phi_j$ is the accuracy rate of the learner's answers to the j_{th} question, using R_j to express the accuracy rate, so $0 \leq R_j \leq 1$. n is the total of the questions which the student replied to test the i_{th} kind of ability. Let $l_i = \text{Ability-how } (A_i)$, and we can obtain the set of learner's CA "Ability-level".

4.3 Estimation Method of Preferences

For the estimation of preferences, it divided into two steps which are initial estimation and dynamic estimation. We have to pretest the learners' interests before them using the system for learning and also initialize the learners' interests. With the proceed of learners' learning activities, we can discover the learners' interests then correct and maintain those values through the data mining about the data of learners' searching concepts, browsing websites' types and topics of discussions. For the initializing values, we can gather users' interest information through users' registration forms or scale, and use direct or indirect matching methods processing those original data to acquire the initial values, as shown in Fig. 4. For dynamic values, it concern about the learners' learning procedure information and historical testing information. Then process the acquired values through the data mining model to obtain the needed data, as Fig. 5 shows.

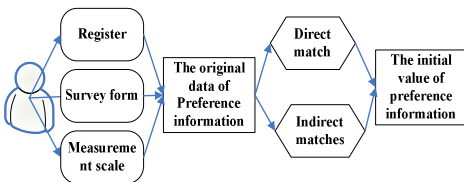


Fig. 4. Initial estimates of preference

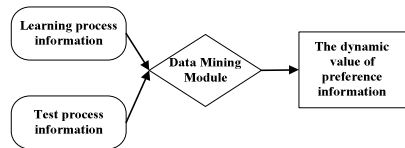


Fig. 5. Dynamic estimates of preference

5 Push Mechanism of Earning Resources in ALM

It selects learning resources which meet the rank needs of each KP of “Knowledge-sequence” in resources base as a collection “R” of learning resources, Then match the “Ability-level” of the learner to the related CA attribute of each resource, Obtain a collection “ R_1 ” of learning resources whose similarity are within a certain range(Ideally, it is necessary to make the recommended learning resources can not exceed any level of CA of the learners, but also help to improve certain CA of learners) Finally, according to the learner’s “preferences-set” selects the learning resources that best meet the learning needs of learners in R_1 . Appropriate for a certain knowledge point learning resources, learners need to be tested after learning a resource. If pass,Then add the mastery of knowledge with 1 pace. Next, determine whether the rank meet the Level requirements of the KP. If it meets, delete the KP .If not, improve the level of KP and continue to learn until meet the demand. If not pass, Then Reselect other resources in this level to learn for learner until he pass the test.

6 Conclusion

The construction of LM is a complex process in Adaptive Learning System. In this paper, based on Gagne’s learning hierarchy theory, We proposed a new LM called Advanced Learner Model (ALM for short) consist of knowledge level, cognitive ability and preferences. We give the representation of ALM separately. Then give acquisition of feature value of ALM in detail. After that, we give push mechanism of learning resources. In addition, this model need to study further in design more precise reasoning methods.

References

1. Wang, X.P., Hu, Q.S., Li, B.: Personalized Information Retrieval for the Internet Intelligent. *Journal of Computer Research and Development* 36, 1039–1046 (1999)
2. Li, Z., Zhao, G.Q.: Research on the Student Modeling of the ICAI System. *Computer Engineering and Science* 3, 101–104 (2002)
3. Zhou, C.P., Zhou, X.Q.: The research of adaptive student model on Bayesian networks. *Journal of Shandong Institute of Light Industry(Natural Science Edition)*, 3–6 (2007)
4. Jia, B., Zhong, S.C., Wang, W., Yang, B.: The Construction and Evolution of Learner Model in Adaptive Learning System. In: Jusoff, H.K. (ed.) *The 2009 International Conference on Computer Technology and Development*, pp. 148–152. IEEE, USA (2009)
5. Lai, X.P., Wu, Y.P., Cheng, W.D.: The Design and Implementation of Multimedia Intelligent CAI in Mining Science. *Computer Systems Applications* 2, 34–35 (1997)
6. Jia, B., Zhong, S.C., Zheng, T.Y., Liu, Z.Y.: The Study and Design of Adaptive Learning System Based on Fuzzy Set Theory. *Transactions on Edutainment*, 1–11 (2010)
7. Gagne, R.M.: Educational technology, as technique. *Educational Technology*, 5–13 (1968)

8. Feng, H., Wang, L., Yang, H.: The Research of Multi-purposes Based Learner Model. *Computer Engineering and Applications* 39, 98–100 (2003)
9. Bloom, B.S., Engelhart, M.D., Furst, E.J.: *Taxonomy of Educational Objectives*, handbook. In: *Cognitive Domain*, London, Longman (1956)
10. Tang, S.Q., Zhong, Z.: Evaluation of Cognition of Students. *Nanning: Journal of Academy of Sciences* 18, 281–284 (2002)
11. Zhuang, Z.: Estimation method of student's cognitive level in self-adaptive distance tutoring system. *Computer Engineering and Applications* 43, 220–223 (2007)

MP-IR: A Market-Oriented Mobile Agents System for Distributed Information Retrieval

Djamel Eddine Menacer¹, Habiba Drias², and Christophe Sibertin-Blanc³

¹ National Computer Science School of Algiers (ESI)

Algiers, Algeria

² Houari Boumedienne Sciences and Technologies University of Algiers (USTHB)

Algiers, Algeria

³ Université Toulouse 1 Sciences Sociales

Toulouse, France

d_menacer@esi.dz, hdrias@usthb.dz, sibertin@irit.fr

Abstract. Most of the Information Retrieval (IR) systems are built on the Client – Server paradigm. While agent-based market systems have relative success, agent-based IR systems seem to fail. Recently, works on a mobile agent-based approach for the search of information on the World Wide Web arise. Indeed, the emergence of mobile agent has given the researchers a new way to achieve efficient mobile agents-based IR systems. This paradigm certainly holds great promise, though the lack of results in reliability and security issues. However, we feel that integrating market mechanisms to a new mobile agent model should improve security while bringing the possibility to solve non market applications problems. In this paper, we present how to use a mobile agent-based approach. The main idea is to generalize the market mechanisms to non-market systems such as IR through an extended mobile agents' model, the seller – buyer model. Then, we present the architecture of a secure mobile agent-based searching system that derives from a general mobile agents-based architecture. Finally, we give an experimental validation to our proposition.

Keywords: Mobile Agent, Security, Information Retrieval, Seller–Buyer model, MP Architecture, Jade, MP-IR.

1 Introduction

The importance of the quality of service (QoS) in distributed applications (non functional aspect) is becoming critical. The quality of service includes (Coulouris et al., 2001): the performance, the security and the safety of functioning (or reliability). While most of the distributed systems are based on the client – server model, new paradigms arise to deal with the growing QoS requirements. Mobile agents are a promising paradigm. The technical aspect of the mobile agents is the main argument for their use (Lange and Oshima, 1999; Rouvrais, 2001). However, the security problem blocks their development (Harrison et al, 1995; Papaioannou, 1999). Indeed, the security problem is *dual*: we have to secure not only the visited hosts but also the visiting agents. The security issues (and solutions) for the hosts protection, known as category

I (Gulyás et al., 2001), are similar to those of traditional systems. However, the main unsolved security problem lies on the possible existence of malicious hosts that can tamper a received agent, known as category III.

A mobile agent needs a specific execution environment on each of the sites that constitute its itinerary. The execution environment is provided by a *mobile agents' platform*. A lot of mobile agent platforms have been developed around the world, such as the Aglets from IBM (Lange, 1998; Lange and Oshima, 1999; ASDK) or the Mole from University of Stuttgart (Baumann et al., 2002). Many efforts are made to provide mobile agents systems that could be more accessible by the industry. FIPA (FIPA, 2009) and MASIF (Milojicic et al., 1998) are the two most known standards. Few mobile agents' platforms provide security features, and none of the existing FIPA and MASIF mobile agents' platforms proposes the protection of the agents.

There are a number of solutions proposed to protect agents against malicious hosts which can be summarized into three lines (Tschudin, 1999; Sander and Tschudin, 1998): make the mobile agents migrating only in a *closed network* and add cryptographic features in order to *detect/prevent agent tampering*.

Recently, some works focus on agent protection (category III) but propose solutions for specific threats (Chan et al., 2000; Garrigues et al., 2008; Garrigues et al., 2009). Finally, most researchers in the area are seeking a better solution, and there is no general methodology suggested to protect agents. In the meantime, developers of mobile agent systems, who are non security experts, have developed their own methodologies according to their own needs.

The remainder of this paper is organized as follows. Section 2 describes some agent-based information retrieval systems. A new mobile agents' model is described in Section 3, and its application through a general mobile agents-based framework is described in Section 4. Section 5 presents how to apply the proposed framework to IR applications and how market mechanisms can perform IR tasks. Section 6 presents the MP-IR platform, a jade adaptation of the proposed framework and describes also an experimental validation. Finally, section 7 concludes the whole paper.

2 Agent-Based Information Retrieval Systems

The current means of locating information on the World Wide Web, known as Web Information Retrieval (IR), relies on the use of *search engines*. These engines attempt to keep up-to-date the information they hold by a variety of means based on spiders, web crawlers ...etc (Hock, 1999). These engines are then queried by users to find information on a particular topic. The use of centralized engines creates bottlenecks when attempting to locate information. The growing size of the information data to be indexed and the processing power required to serve search requests will affect the ability of the search engines' technology to continue to cope. At any time, any given search engine is estimated to cover no more than 40% of the web in its database. To perform an exhaustive search requires the user to employ several search engines and to assume that each one has access to a different 40%. To remove the bottleneck problem, the index needs to be distributed. The growth of interest in mobile agent systems leads the authors in (Grey et al., 2007) to suggest that one possible means for achieving this end is to use mobile agents which wander through the web (in a directed

fashion) seeking the information on behalf of the users. The proposed scheme (AgentSeek system) involves three separate types of mobile agent: “ferrets” which act on behalf of web searcher users, seek for information providers and advertise the location of information consumers, “publicists” which act on behalf of web site creators (people providing information), advertise the location of information providers and seek information consumers, and “gurus” which facilitate encounters between ferrets and publicists. Beside the fact that the proposed scheme uses specific concepts, ferrets, gurus and publicists that cannot be applied to other systems and the security aspect had not been taken into account in this study.

In (Arcangeli et al, 2006), an interesting study proposes an approach based on mobile agents and components in order to simplify the development and the deployment of adaptable information retrieval systems in the context of distributed heterogeneous peer-to-peer networks. This promising proposal uses mobile agents as a solution to the deployment problem. The security aspect is discussed as a perspective for the system.

Some IR systems also exist in the industry. NetSA (Coté et al., 2002) is a multi-agents system for the IR on heterogeneous distributed sources. This system comprises essentially the following agents: *User agents* that collect and filter information from and to the clients; the *Broker agents* that associate the requests to agents which are able to respond to them; the *Resources agents*, which are linked to an information resource (internal or external) and are able to update the data. However, the NetSA system is a static multi-agent system that does not implement the agents’ security. Calvin (Bauer and Leak, 2002) is a secured multi-agents system that provides the following agents: *Calvin Web, an interface agent; Analysis agents (TFIDF, WordSieve and DocStats) that perform analysis of the user’s profile and behaviour and Research Agents (AltaBot and GoogleBot) that perform profile-based searches for the user.* The Calvin system is also a static multi-agent system. Calvin implements a security level in multi-agents system by authentication and encrypted communications between the agents.

Our approach should also use distributed indexes, user’s agents, provider’s agents and facilitator’s agents. However, the originality of our approach resides in the fact that it relies upon the generalization of market mechanisms to non-market systems. The success of the agents in e-commerce systems gives us the opportunity to use these well-experienced concepts in IR systems. Another difference, and not the least, is that our approach is based on a software *secure* mobile agents model. While most of mobile agents’ platforms provide asymmetric security solution (Gulyás et al., 2001), we propose a symmetric security model.

3 The Seller – Buyer Model

Our first contribution is to propose to extend the classical Mobile Agent (MA) model towards a more suitable model, the Seller-Buyer model (SB) (Menacer et al., 2009). The SB model allows to keep at least the same performance level as the Mobile Agents model (figure 1), and to reduce the mobile agents security problem. This is done by preventing an agent from migrating *directly* to a host and a host from receiving mobile agents. This is possible if we dissociate the rendering of services and the hosting of visiting agents; then a mobile agent representing the *client* meets, on *intermediate sites*, a service *provider* representative. As the client agent, the server agent

process becomes mobile. The client agents are *delegated* by clients to *buy* a service (buyer agents) and the server agents are *delegated* by the remote services to *sell* a service (seller agents). Mobile agents (buyer and seller) may migrate *only* on meeting sites: the *market places* or *MP*. In the SB model, we adapt market mechanisms to distributed systems. The seller agents are then in *competition*.

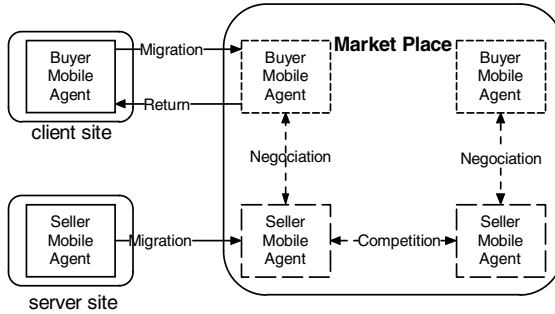


Fig. 1. The seller-buyer model

Security. While security issues of mobile agents are *symmetric* (they concern the protection of both the agents and the hosts), in contrast to this, the research efforts concerning these issues seem to be *asymmetric* (focus on host security). The SB model reduces the security problem by a symmetric approach: *only trusted agents (cat. I) can migrate on trusted MP (cat. III)*. This is shown in table 1.

Table 1. SB security features

Category	Feature	Description
I	Certified and signed mobile agents	Only these trusted agents can migrate on the MP
III	Closed network	Agents migrate only on the MP
	Agent tampering prevention	The MPs are certified and high trusted hosts. The only permitted migration is the migration on the these trusted MP
	Agent tampering detection	Certified agents are <i>digitally signed</i> . Recovering from attacks is lesser in a closed network

4 The SB Model for Distributed Applications: The MP Architecture

We propose the *MP (Market - Place) architecture* based upon the SB model for the development of mobile agents distributed applications. This architecture can be considered as a closed space in which the agents, representing the clients and the providers, move to search or offer services.

4.1 MP Components

There are two *external* actors to the MP architecture: *clients* that send requests by means of buyer mobile agents and *providers* that offer services at MP sites by means of seller mobile agents. The novelty of our architecture is to propose a general framework that can deal with any type of distributed application. The basic idea is that each service S in MP architecture belongs to a *class of services* SC . A class of services SC belongs to an *application domain* D . We can put: $D = \{SC_1, SC_2, \dots, SC_N\}$ and $SC_i = \{S_{i1}, S_{i2}, \dots, S_{iM}\}$

The MP architecture comprises the following components:

The Market Places (MP). In order to give mobile agents a *directed way* to request a service, the MP sites are organized according to the *class* of the offered services. One MP hosts one service class SC_i (so one or several services ($S_{i1}, S_{i2}, \dots, S_{iN}$) belonging to the class SC_i), but one service class can be hosted by several MP. The services are located, within the MP site, in *e-shops* (Wang et al., 2002). In the same manner, each e-shop hosts one service, but one service can be hosted by several e-shops. Therefore, seller agents visit e-shops and the dialogue between buyer agents and seller agents takes place in e-shops. The market place holds a special database service called MP Directory Service (MPDS) that manages information about e-shops and agents present in the place. MP sites are identified by *IP* address or URL.

The Agent Service Providers (ASP). The client and provider users need a mean to create their mobile agents. A user connects via its browser to a specific site. This site is responsible for the creation of mobile buyers or seller agents, according to the type of user (client or provider). We refer this site to as Agent Service Provider (*ASP*) (Wang et al., 2002). An ASP is accessible by a URL, for example mp://www.it.com.

The MP Name Servers (MPNS). When a mobile agent requests or offers a service S that belongs to the class SC , it searches MP sites providing the class SC . To do this, the agent sends a request to MP Name Servers (MPNS), analogue to the DNS servers on the Internet. The answer is a list of IP addresses of MP sites that offer the class of service SC , constituting the *itinerary* of the agent: (mp_1, mp_2, \dots, mp_N).

The Trust and Security Authorities (TSA). A mobile agent must be *certified* before it visits MPs. We propose to add *PKI* (Public Key Infrastructure) components to the architecture. The ASP provides a pair of keys (private, public) to mobile agents by means of a cryptography service. In the case of a buyer agent, the certification authority delivers a *user certificate* linked to the public key of the agent; in the case of a seller agent, a *server certificate* is issued. The mobile agent can be signed or encrypted. We propose to host the *certification authorities* in special sites called Trust & Security Authority (*TSA*). All the components of the MP architecture should be certified: ASP, MP, MPNS and the TSA themselves.

Agents in MP Architecture. There are two types of agents: *mobile agents* that comprise buyer agents and seller agents, and *static agents* that comprise manager agents and facilitator's agents. Manager agents manage different sites in the system: ASP, MP and the e-shops. For MP site, the manager agent is called MPSPM (MP Service Manager) Agent. The facilitator agents are responsible for recording information

about mobile agents running on their place or site, and also for providing migration services for mobile agents running on other places or sites. There is one facilitator agent at the following hosts: ASP, MP and e-shops.

4.2 The MP Architecture Diagram

Figure 2 shows the MP architecture diagram (in **bold** the agents' migrations):

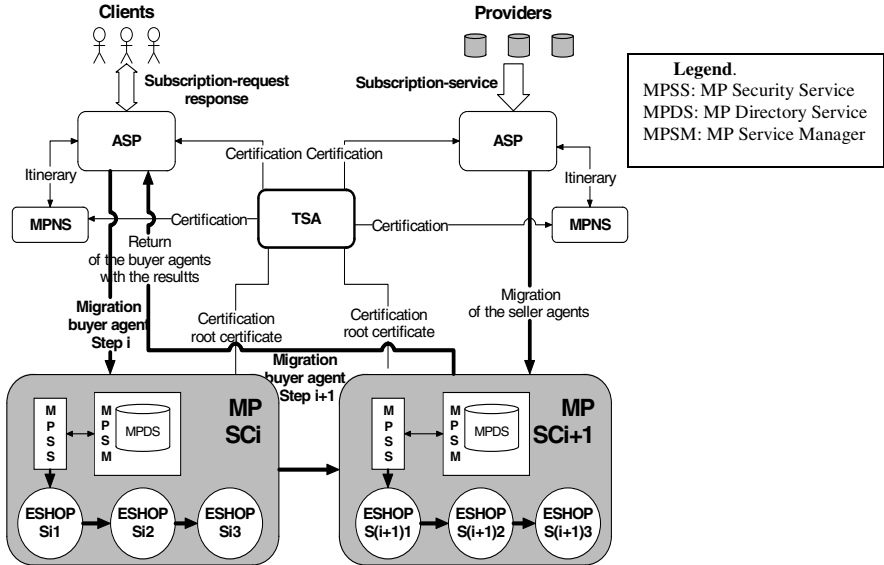


Fig. 2. MP system architecture

5 Application to the Distributed Information Retrieval (DIR)

We define the application domain "IR" as including the class of service as "Search Category", and the service as "Search theme". For example: D=IR, SC=general, S=general, for general purpose search; D=IR, CS=IT, S=software, for specific search.

The MPs are classified according to the *search category* they offer. A buyer mobile agent is a *search agent* (also called *metasearch agent*) which acts on behalf of a user. A seller agent owns an index and a search code, and can be considered as *search engines*. The search engine agent acts on behalf of a provider. It is reasonable that a seller agent carries only an index of a search theme S. To facilitate the users' searches, we propose a *distributed index* through search engine agents.

According to the search category SC of the request, the asked MPNS server gives the metasearch agent an itinerary composed of the MPs belonging to this search category. Finally, the metasearch agent meets the search engine agents in the e-shop corresponding to the search theme S and located in a MP site that belongs to the search category SC; the metasearch agent can then ask several search engine agents, filter the different results and return the best result to the user. Figure 3 shows interaction between a metasearch and search engine agents within a MP.

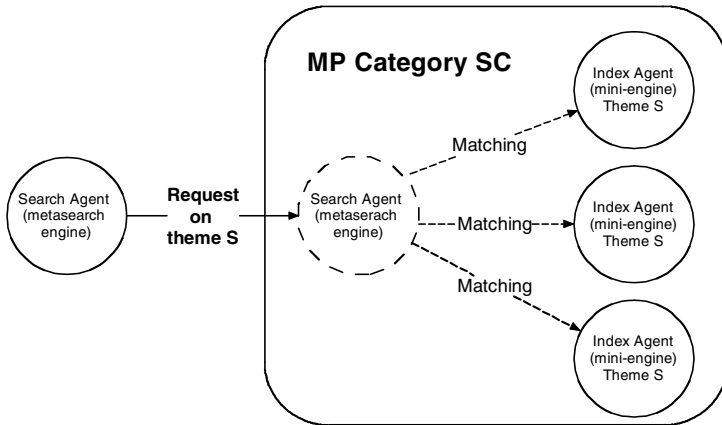


Fig. 3. Interaction between metasearch and engine agents

E-shops can be considered as negotiation rooms. The negotiation process must be adapted to the Information Retrieval. Several search engine agents can offer the same service *S* but with different *qualities of service (QoS)*. The QoS may be evaluated according to the *relevance* and the *size* of returned results. We assume that each search engine agent is able to return, in addition of the results for the request, the average *relevance* and the *size* of these results. The negotiation process based on the relevance *R* and the size of results *S* can be summarized in the algorithm 1:

```

For each MP in the metasearch agent's itinerary
  For each eshop in the MP
    /* CFP process */
    1. The metasearch agent makes a call of proposal by sending
       a CFP message (SC, S) to all search engine agents
       present in the eshop;
    2. The search engine agents interested by the CFP (selected)
       answer the metasearch agent by sending a service offer;
       /* Ns is the number of selected agents */
    /* Reverse auction process */
    3. The initiator of the auction, the metasearch agent,
       defines the wished (and hidden) price that reflects
       the Rmin relevance and the Szmax size of the results
       corresponding to the search request.
    4. The metasearch agent sends its request (SC, S, (k1,
       k2, ..., kn)) to the selected search engine agents
       (matching); /* ki are the keywords */
    5. For each round in (1.. Rdmax) /* Rdmax is the maximum
       number of rounds allowed */
       a. While number of iterations j < Jmin,
          (1 < Jmin ≤ Ns) /* Jmin is the minimum number of
             iterations */
             Each selected search engine j (0 ≤ j ≤ Jmin-1)
             sends public proposition (SC, S, Rj, Szj) to the
             metasearch agent;
           End while;

```

```

    b. if ( $R_j < R$  or  $S_j > S_z \forall j$ ) then
        Search engine agents are invited to decrease their propositions for another round (decrease  $S_z$  and/or increase  $R$ );
        Else the metasearch agent selects the three most suitable propositions (the answers that feature the maximum relevance  $R$ , the minimum size  $S_z$  and the auction ends.
        Exit; /* End of auction */
        End if;
Next round;
6. The metasearch stores the results ( $SC, S, (url_1, url_2, \dots, url_m), S_z, R$ ) in its memory; /*  $url_i$  are the URL of the relevant documents */
Move to next eshop;
Move to next place;

```

Algorithm 1. Interactions between metasearch and search agents

In the ASP: Merging, filtering and clustering the results according to the client *profile*. As a result, we have defined the outlines of MP-based Information Retrieval architecture. We refer this architecture to as the *MP-IR*.

6 A Jade Implementation of MP-IR

6.1 Implementation Overview

Jade (for Java Agent DEvelopment Framework) (Tilab, 2010; Bellifemine et al., 2001; Fortino et al., 2010) is a free and open source platform for the development of FIPA agents-based systems. The MP architecture can be implemented as a set of Jade platforms distributed over several computers in a network. A *Market place* is a set of computers including a jade main platform server and one or several jade platforms (containers) servers that implement e-shops. The ASP site is a main Jade platform which allows the creation of mobile agents in the system. The other MP components (MPNS and TSA) may be java services. The communication between Jade platforms and these servers are performed by sockets.

6.2 Agents in MP-IR

Every agent inherits the *Agent* class of the package *jade.core.agent*. The tasks of each Jade agent are called *behaviours*. Jade allocates one thread for each agent.

Static Agents. These agents provide services to the system. The DF agent of a MP acts as a MP facilitator agent. The AMS agent of the main platform manages the MPs and acts as the MPSM agent. There is only one DF in a jade platform. The static agents implement a *cyclicBehaviour* (a repetitive behaviour issued from the class *CyclicBehaviour* of the package *jade.core.behaviours*) since they run repetitive tasks.

Mobile Agents. The *metasearch agents* may have an advanced decision autonomy model. We think that the Jade finite state machine (FSM) model is suitable for this type of agent. The FSM are instances of the class *FSMBehaviour* of the package *jade.core.behaviours.FSMBehaviour* and can implement behaviours. The *request* of the client is included in the behaviour.

The *search engine* agents may also implement a finite state machine but lighter than those of the metasearch agents because it does not performs multiple migrations as metasearch agents do. The *service (code+index)* of the provider is included in the behaviour. At its arrival in a MP, a search engine agent registers its services to the sub-DF agent of the appropriate e-shop. Finally, a search engine agent interacts remotely with its provider site by sockets.

6.3 Experimental Validation

The quality of service in IR systems can be measured by the response time. However, the users are more concerned by the quality of the results rather than the time response. The quality of the results is measured by the relevance. The relevance can be measured by two factors: the precision and the recall. The recall measures the ability of the system to retrieve all relevant documents. The precision measures the ability of the system to retrieve only relevant documents or reject all irrelevant documents.

We did our tests on a specific benchmark called a corpus. The corpus contains a set of documents, a set of requests and a list of relevant documents for each request. The corpus we used contains 500 documents and 40 requests. We have developed two IR systems. The first one is a classical IR system using client –server interactions and the second one is MP-based IR system using market interactions according to the algorithm 1. The classical IR system is based upon Terrier (Terrier). Both systems use the same corpus. The measures we did are based upon the recall/precision curves. The table 2 summarizes the average precision depending of the recall for the 40 requests.

Table 2. Summary of results

	Standard recall levels (%)										
	0	10	20	30	40	50	60	70	80	90	100
Terrier	69.96	55.45	48.67	31.48	22.07	16.03	12.65	10.88	9.12	5.55	1.75
MP-IR	60.51	50.44	43.92	29.17	20.05	15.00	11.64	10.88	9.12	5.55	1.75

The figure 4 shows the average precision curve using the 11 standard recall levels for the 40 requests. Globally, it is interesting to note that, although Terrier is better, both curves are almost similar.

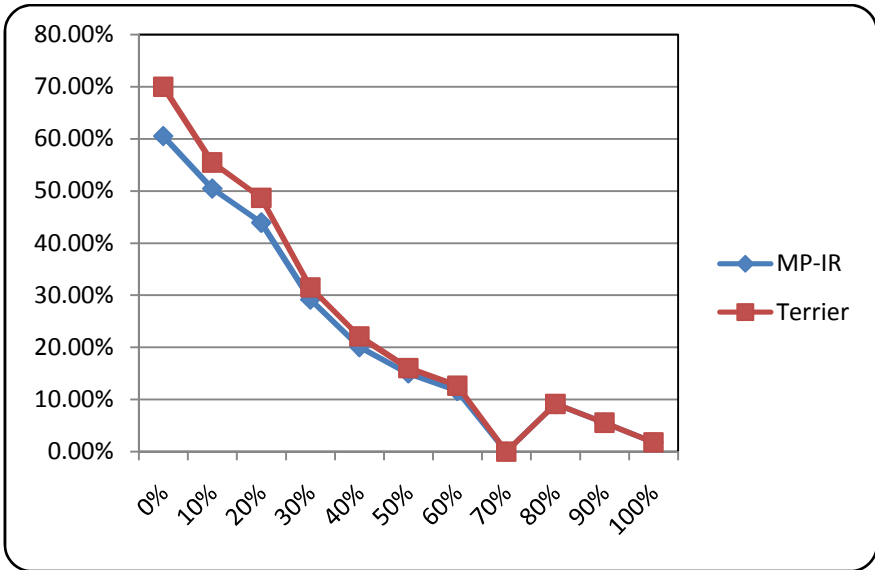


Fig. 4. Average precision of both systems

7 Conclusion and Future Works

The success of the mobile agents' paradigm for building distributed systems for the IR depends upon the agents' security. On the Internet, the Web IR requires more security and interoperability between the different platforms. While most of the related works focus on specific solutions, our approach is more general. This is done by a generalization of the market mechanisms to non-market systems. The classical mobile agents' model is not suitable for building secure distributed systems on large-scale networks such as the Internet. Therefore, we have extended this model by adding market mechanisms. While most of mobile agents' platforms provide asymmetric security solution, the extended model, the seller – buyer model provides symmetric security mechanisms. Based upon this new model, we develop a global architectural design called MP architecture. Finally, we apply MP architecture to IR systems by introducing an interaction IR model between metasearch agents and search engine agents. We feel that this model improves the IR process. We project therefore to complete the IR prototype and to perform further experiments against classical IR systems.

References

1. Chan, A.H.W., Wong, C.K.M., Wong, T.Y., Lyu, M.R.: Securing Mobile Agents for Electronic Commerce: an Experiment. In: Proceedings of the IFIP TC11 Fifteenth Annual Working Conference on Information Security for Global Information Infrastructures. Kluwer, B.V. Deventer (2000)

2. Arcangeli, J.P., Leriche, S., Pantel, M.: Une infrastructure distribuée flexible à agents et à composants. Journée ASR / GSP / ADAPT, ENST, Paris, 6 Avril 2006 (2006)
3. ASDK. IBM Aglets Software Development Kit Homepage, <http://www.trl.ibm.co.jp/aglets> (accessed in July 2010)
4. Bauer, T., Leak, D.B.: Handling Complex Information Environments: A Multi-Agent Framework. In: Giorgini, P., Lespérance, Y., Wagner, G., Yu, E.S.K. (eds.) Proceedings of the Fourth International Bi-Conference Workshop on Agent-Oriented Information Systems (AOIS-2002 at AAMAS 2002), Bologna, Italy. CEUR Workshop Proceedings, vol. 59. CEUR-WS.org (2002)
5. Baumann, J., Hohl, F., Rothermel, K., Straber, M., Theilmann, W.: Mole: A mobile agent system. *Journal of Software-Practice and Experience (SPE)* 32(6), 575–603 (2002)
6. Bellifemine, F., Poggi, A., Rimassa, G.: Developing multi-agent systems with a FIPA compliant agent framework. *Journal of Software Practice and Experience* 31(2), 103–128 (2001)
7. Coté, M., Chaib-Draa, B., Troudi, N.: NetSA: A reusable multi-agents architecture for rich information environments. Cépadués-Edition (2002)
8. Coulouris, G.F., Dollimore, J., Kindberg, T.: *Distributed Systems*. In: Concepts and Design, 3rd edn. Addison-Wesley (2001)
9. FIPA. Official site of the Foundation for Intelligent Physical Agents. Obtained through the Internet (2009), <http://www.fipa.org> (accessed 15/06/2009)
10. Fortino, G., Garro, A., Mascillaro, S., Russo, W.: Using event-driven lightweight DSC-based agents for MAS modeling. *International Journal of Agent-Oriented Software Engineering* 4(2), 113–140 (2010)
11. Garrigues, C., Robles, S., Borrell, J.: Securing dynamic itineraries for mobile agent applications. *Journal of Network & Computer Applications* 31(4), 487–508 (2008)
12. Garrigues, C., Midas, N., Buchanan, W., Robles, S., Borrell, J.: Protecting mobile agents from external replay attacks. *Journal of System Software* 82(2), 197–206 (2009)
13. Grey, D.J., Dunne, G., Ferguson, R.I.: A Mobile Agent Architecture for Searching the WWW. Scientific Commons (2007), paper available under, <http://en.scientificcommons.org/42430563>
14. Gulyás, L., Kovács, L., Micsik, A., Pataki, B., Zsámboki, I.: An Overview of Mobile Software Systems. MTA SZTAKI Technical Report. Department of Distributed Systems, MTA SZTAKI, Computer and Automation Research Institute of the Hungarian Academy of Sciences (2001)
15. Harrison, C.G., Chess, D.M., Kershenbaum, A.: Mobile Agents: Are they a good idea? IBM T. J. Watson Research Report. RC 19887 (1995)
16. Hock, R.: Web search engines: Features and commands. *Search Engine Section* 23(3), 24–28 (1999)
17. Lange, D.B.: Present and Future Trends of Mobile Agent Technology. In: Rothermel, K., Hohl, F. (eds.) MA 1998. LNCS, vol. 1477, p. 1. Springer, Heidelberg (1998)
18. Lange, D.B., Oshima, M.: Seven Good Reasons for Mobile Agents. *Communications of the ACM* 42(3), 88–89 (1999)
19. Menacer, D.E., Drias, H., Sibertin-Blanc, C.: The Seller – Buyer model: an extension of the Client – Server model by the mobile agents. In: Rodrigues, L., Barbosa, P. (eds.) Proceeding of the IADIS International Conference Intelligent Systems and Agents 2009, Algarve, Portugal, June 21–23, pp. 133–140 (2009) ISBN: 978-972-8924-87-4

20. Milojcic, D., Breugst, M., Busse, I., Campbell, J., Covaci, S., Friedman, B., Kosaka, K., Lange, D., Ono, K., Oshima, M., Tham, C., Virdhagriswaran, S., White, J.: MASIF: The OMG Mobile Agent System Interoperability Facility. In: Rothermel, K., Hohl, F. (eds.) MA 1998. LNCS, vol. 1477, pp. 50–67. Springer, Heidelberg (1998)
21. Papaioannou, T.: Mobile Agents: Are They Useful for Establishing a Virtual Presence in Space? In: Agents with Adjustable Autonomy Symposium. part of the AAAI 1999 Spring Symposium Series (1999)
22. Rouvrais, S.: Construction de services distribués: une approche à base d’agents mobiles. TSI-HERMES / RSTI – TSI 21(7), 985–1007 (2002)
23. Sander, T., Tschudin, C.F.: Protecting Mobile Agents Against Malicious Hosts. In: Vigna, G. (ed.) Mobile Agents and Security. LNCS, vol. 1419, pp. 44–60. Springer, Heidelberg (1998)
24. Terrier, <http://www.terrier.org> (accessed in October 2010)
25. Tilab. Jade official site, <http://jade.tilab.com> (accessed in July 2010)
26. Tschudin, C.: Mobile Agent Security. In: Intelligent Information Agent: Agent Based Information Discovery and Management in the Internet, pp. 431–446 (1999)
27. Wang, Y., Tan, K.L., Ren, J.: A Study of Building Internet Marketplaces on the Basis of Mobile Agents for Parallel Processing. Journal of World Wide Web 5(1), 41–66 (2002)

Framework for Goal-Driven Negotiation Process

Ying Lei

School of Economics and Management, Henan Polytechnic University,
Jiaozuo, China

lei1978ying@yahoo.com.cn

Abstract. Agent-based negotiation support has been applied in electronic commerce and provides users with valuable information about products and services. In this paper a framework for goal-driven negotiation process is proposed which can be used to design negotiating agents. The framework includes goal representation, goal selection, goal execution and goal reconsideration. Goals are set to control agents' inference processes and determine negotiation actions. Instead of focusing on predicting possible negotiation outcomes, the framework emphasizes rational and explainable deliberation process of negotiating agents.

Keywords: negotiation support, agents, goals.

1 Introduction

Researches on agent-mediated electronic commerce have become very popular in recent years and agents can perform negotiation tasks instead of human users in certain contexts. In such processes agents should "understand" their user's requirements and try their best to satisfy them. Owing to the complexity of negotiations how to design negotiation support system becomes very tough. Many researchers provide their suggestions in various angles and hope agents can greatly improve negotiation efficiency and help users to make agreements in shorter time. There have emerged many experimental systems such as AuctionBot, Aspire and these systems have demonstrated the great potentials of agent technologies in electronic commerce domain [1, 2]. Now most of negotiation support systems focus on predicting negotiation outcomes and provide possible agreements based on information from users. Agents should collaborate with users to reach satisfying and explainable agreements during negotiations.

Negotiating agents elicit information from their users including preferences, goals, strategies, etc. Negotiation goals are abstract representation of negotiator's demands, e.g., a product specification with a price to match. Preferences can be considered as constraints which can help agents make negotiation decisions. While negotiation strategies are concrete plans which can be used to achieve negotiation goals, they can either be rule-based/argumentative, game-theoretic or heuristic/adaptive [3]. Most researches on agent-based negotiation concentrate on negotiation protocols which governing the negotiation processes and negotiation strategies which being used to determine the exact negotiation activities [4-7]. The roles of goals have been neglected and just replaced by simple values of negotiation issues such as reservation

prices. Although both negotiation protocols and strategies are very important for designing negotiation support system, from a different point of view goals can play more important role in such processes and negotiation processes can be considered as a goal-seeking and goal-execution process of agents.

Although design of agent-based negotiation support has been investigated by researchers both a theoretical and a practical perspective [8], there is great gap between these researches. On the one hand theoretical researches mainly attempt to develop formal models of negotiation and these models are often based on game-theoretical techniques, rich but restrictive. Assumptions of these models limit their applicability to solve real negotiation problems. On the other hand practical application researches only adopt very simple agent models which cover very limited negotiation scenarios (e.g. auctions). Activated by agents' following characteristics we emphasize the role of goals in agent-based negotiation support.

- Reactive: agents are able to perceive their environment and respond in a timely fashion to changes within it in order to satisfy their goals.
- Proactive: agents are able to exhibit goal-directed actions by taking the initiative in order to satisfy their goals.
- Social ability: agents are capable of interacting with other agents (or humans) in order to satisfy their goals.
- Autonomy: agents operate without the direct intervention of humans or others, and have some kind of control over their actions and internal states.

The remainder of this paper is structured as follows. The next section describes the life cycle of goal-driven negotiation process. Section 3 analyzes the different stages of such process including goal representation, goal selection, goal execution and goal re-consideration. The final section provides some conclusions and highlights open issues that need to be more fully addressed.

2 Goal-Driven View of Negotiation Process

The motivational power of goals drives agents' actions and agents must have the capabilities to reason about them. Goals specifying what must be achieved without specifying how, and in that sense, enable agents to choose the best means available to them in deciding how to achieve them. However, it is necessary to analyze how certain goals are generated. Researchers propose that motivations behind goals address these both by providing the reasons for goals and by offering constraints on how the goals might best be achieved when faced with alternative courses of actions. In that sense, motivations both release and constrain agents' autonomy [9]. Analyzing motivations of negotiators also provide the base to realize interest-based negotiation rather than position-based negotiation and understanding motivations behind goals facilitates integrative negotiation which may result in win-win negotiation outcomes. Goals may have hierarchical structure and these goal levels are internal to agents. At these levels the aspirations, reservation values, trade-offs and preferences are considered [10]. Agents need to reason and decide by themselves how to select goals and what actions they should take so that their goals can be achieved successfully. Agents should also be able to improve their behaviors over time, that is, they may become

better with experiences at selecting goals and achieving goals by taking correct actions [11].

When facing negotiation problems, negotiators may have various goals. It is very important for agents to find out what negotiators want and why they want them. Negotiation goals are influenced by negotiation contexts greatly and the same negotiation problem in different contexts may result in different goals. Agents negotiate on behalf of their users and they first elicit information about users' requirements, preferences, motivations, etc and analyze negotiation contexts. These information are used to form agents' high-level goals. Secondly, agents should decompose these goals and select appropriate goals which pursued currently based on negotiation states, goal priorities and other information. Thirdly, actions are taken by agents towards goals. But this is usually not enough since actions may not result in the anticipated outcomes. Obviously, the practical outcomes are influenced by many factors, most important of which are what other agents do. It is necessary to evaluate outcomes of actions according to goals, check if goals have been achieved and adjust actions according to the evaluation results, and this in turn may result in the reconsideration of goals. The life cycle of goal-driven negotiation process comprises following stages: negotiation problem definition, goal representation, goal selection, goal execution and goal reconsideration (see Fig. 1).

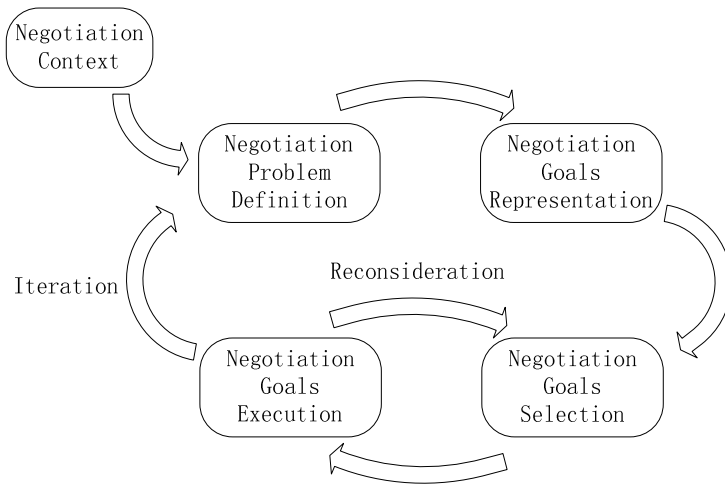


Fig. 1. Life Cycle of Goal-Driven Negotiation Process

The whole process begins with negotiation problem definition when negotiation parties specify the requirements of negotiation. Then users delegate their “initial” goals to agents and agents select goals which are pursued and execute these goals. The overall process often requires iteration and reconsideration. Iteration refers to repeating the negotiation processes when necessary such as when negotiations break down agents must renegotiate. Reconsideration may include decomposing high-level goals into subgoals again, dropping some goals or relaxing goal constraints.

3 Framework for Goal-Driven Negotiation Process

According to the life cycle of goal-driven negotiation process, the first stage is negotiation problem definition. Most of negotiations include multiple issues, for example price, delivery time and quality. Negotiation parties should specify all possible issues of negotiation objects. In addition, other information about negotiation contexts should also be collected and processed.

3.1 Goal Delegation and Goal Representation

When users delegate their negotiation tasks to agents, all efforts will be made to ease users' work. First, agents should embrace the motivations of users and try to achieve goals derived from these motivations. Secondly, users need only provide declarative goals instead of explicit operational strategies, that is, users are interested only in the negotiation outcomes and not in the specific way agents achieve them. Certainly, if necessary, agents can explain their negotiation processes in order to improve users' trust in them, and users should believe that agents know how to achieve these goals.

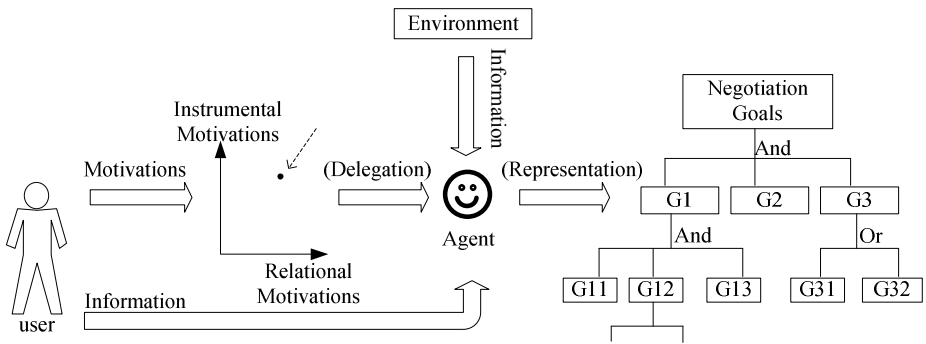


Fig. 2. Process of Goal Delegation and Goal Representation

In goal delegation process motivations are used to guide agents' goal-setting processes. Motivations can be considered as users' high-order desires and preferences. Users' negotiation motivations may include two aspects at least: generally, users must have multiple instrumental motivations, such as obtaining money, goods, services, or information; on the other hand relational motivations are mainly concerned with gaining power and building a level of trust [12]. If users only concern with short-term relationship, relational motivations may be relatively unimportant. While if users want to build long-term relationship, relational motivations are likely to play a more important role on how negotiations are performed. Each of agents' motivations derived from users' motivations can be characterized by intensity. When a motivation has high intensity, any goals related that motivation should be more important. Thus, when agents face a large number of goals, they can quickly determine their priorities and proceed to organize their reasoning and actions accordingly. Especially, agents can reconsider their goals based on motivations. Different motivations

form a multi-dimensional space and goals are situated in different positions of the space. According to the intensity values of different motivations, relevant goals can be activated by agents during negotiation (see Fig. 2).

Most of negotiations are not merely content to achieve a single goal and agents must achieve multiple goals. Typically, goals have hierarchical relations and can be represented as a tree graph. Relations between different goals are complex. Traditional goal analysis decomposes goals into subgoals by using AND/OR-relations. If goals are decomposed by using AND-(OR-) relation, all subgoals (at least one subgoal) should be satisfied. Topology structure of goals involves AND/OR-relations at least. However, there have more complex relations between goals of agents as shown in Fig. 3.

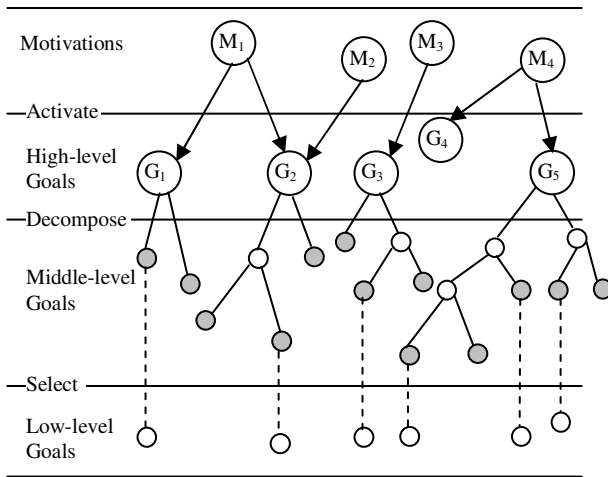


Fig. 3. Topology Structure of Goals

Fig. 3 shows that high-level goals are activated by motivations and can be further decomposed into executable goals (grey nodes in Fig. 3) which directly defined on concrete negotiation issues. Low-level goals are selected from the set of executable goals based on some rules and they form the current pursued goals of agents. Executable goals correspond to middle-level goals that cannot be decomposed further and there exist very complex relations among various middle-level goals. Six types of relations are shown in Fig. 4.

- (1) As shown in figure (a), there exists AND-relation between subgoals g_{11}, g_{12} , that is, the satisfaction of goal g_1 needs both g_{11} and g_{12} are satisfied.
- (2) As shown in figure (b), there exists OR-relation between subgoals g_{11}, g_{12} , that is, the satisfaction of goal g_1 needs at least one of the subgoals g_{11}, g_{12} is satisfied.
- (3) As shown in figure (c), there exists partial support (+) relation between goal g_i and goal g_j . Namely, the support degree between g_i and g_j is greater than

the interference degree between them on negotiation space which is composed of various values of negotiation issues.

- (4) As shown in figure (d), there exists partial interference (-) relation between goal g_i and goal g_j . Namely, the interference degree between g_i and g_j is greater than the support degree between them on negotiation space.
- (5) As shown in figure (e), there exists full support (++) relation between goal g_i and goal g_j . That is, goal g_j only supports goal g_i and doesn't interfere it.
- (6) As shown in figure (f), there exists full interference (--) relation between goal g_i and goal g_j .

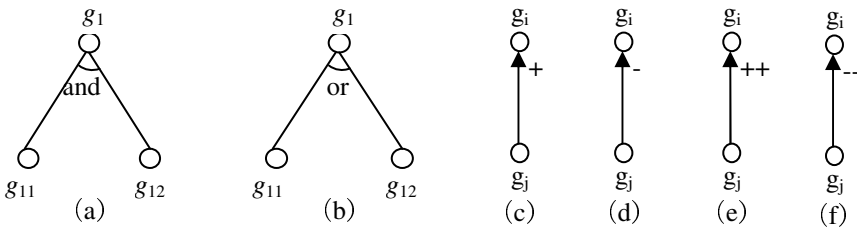


Fig. 4. Relations between Middle-level Goals

As shown in Fig. 4, AND/OR-relations form the hierarchical structure of goal graph and they constitute various AND/OR trees. However, the four latter relations (shown in figure (c)-(d)) specify the relations between executable goals of different AND/OR trees. And if there haven't any support or interference relationships between executable goals, they are independent.

3.2 Goal Selection

Goals do not possess truth conditions, but conditions of satisfaction. Goals are either achieved or not achieved, or achieved to various degrees. In addition, some goals currently may not be pursued in favor of other goals, and others currently try to achieve. Goals can be in different processing states and relations between goals can greatly influence negotiation outcomes. There are two problems should be handled carefully. The first is how agents select low-level goals. Agents should have capabilities to determine their pursued goals according to negotiation states. Secondly, it is of importance which negotiation strategy agents can use to achieve their goals. Negotiation strategy selection and execution are pivotal contents which require further researches, such as how strategies can be determined, how agents can recover from strategy failures and try other methods to achieve goals. Changes in negotiation states, goal priorities and beliefs about the possibility or otherwise of achieving a given goal mean that the set of goals selected may change over time.

We provide an approach to analyze the relations between goals. Conflicts between different goals are mainly due to interdependence and it is impossible to satisfy these goals simultaneously. Agents should analyze the relations between goals and take actions to satisfy all goals as much as possible. The possible values of negotiation issues

constitute negotiation space of agents and provide various potential negotiation offers. Suppose the negotiation space is O , the set of low-level goals is $\{g_1, g_2, \dots, g_k\}$ and $o \in O$ is a negotiation offer. Then we consider following problem:

$$CG(o) = \underset{o \in O}{Opt}(sat_{g_1}(o), \dots, sat_{g_k}(o))$$

where $sat_{g_i}(o)$ is function used to compute the satisfaction degree of goal g_i when negotiation offer is o , $sat_{g_i}(o) \in [0,1]$. Agents should try to achieve negotiation outcomes that have optimal satisfaction degrees of goals. $CG(\cdot)$ can be defined as a function to compute such optimal values, or a set of rules to evaluate outcomes. Next, we can define the support and interference relations between goals.

Definition 1 (support relation between goals): $\forall o_1, o_2 \in O$, if $sat_{g_i}(o_1) > sat_{g_i}(o_2)$, then $sat_{g_j}(o_1) > sat_{g_j}(o_2)$. There exists support relation between g_i and g_j .

Definition 2 (interference relation between goals): $\forall o_1, o_2 \in O$, if $sat_{g_i}(o_1) > sat_{g_i}(o_2)$ then $sat_{g_j}(o_1) < sat_{g_j}(o_2)$. There exists interference relation between g_i and g_j .

In other situations goals are independent. Based on the support and interference relations we can classify the negotiation space O into three subspaces O^+ , O^- and $O^\#$ ($O = O^+ \cup O^- \cup O^\#$, $O^+ \cap O^- = \emptyset$, $O^+ \cap O^\# = \emptyset$, $O^- \cap O^\# = \emptyset$), where O^+ denotes the subspace in which there exist support relations between goals; O^- denotes the subspace in which there exist interference relations between goals; while in subspace $O^\#$ goals are independent. Then we can compute the interdependence degree of goals on negotiation space.

Definition 3 (support degree between goals): support degree between goal g_i and goal g_j can be computed as follows:

$$sd(g_i, g_j) = \frac{\sum_{o_1, o_2 \in O^+} (|sat_{g_i}(o_1) - sat_{g_i}(o_2)| + |sat_{g_j}(o_1) - sat_{g_j}(o_2)|)}{\sum_{o_3, o_4 \in O} (|sat_{g_i}(o_3) - sat_{g_i}(o_4)| + |sat_{g_j}(o_3) - sat_{g_j}(o_4)|)}$$

Definition 4 (interference degree between goals): interference degree between goal g_i and goal g_j can be computed as follows:

$$cd(g_i, g_j) = \frac{\sum_{o_1, o_2 \in O^-} (|sat_{g_i}(o_1) - sat_{g_i}(o_2)| + |sat_{g_j}(o_1) - sat_{g_j}(o_2)|)}{\sum_{o_3, o_4 \in O} (|sat_{g_i}(o_3) - sat_{g_i}(o_4)| + |sat_{g_j}(o_3) - sat_{g_j}(o_4)|)}$$

We can determine the support and interference degree between goals based on the satisfaction degrees of various goals. If $sd(g_i, g_j) > cd(g_i, g_j)$, then there exists partial

support (+) relation between goals. If $sd(g_i, g_j) < cd(g_i, g_j)$, then there exists partial interference(-) relation between goals. If $sd(g_i, g_j) > 0$ and $cd(g_i, g_j) = 0$, then there exists full support (++) relation between goals. If $sd(g_i, g_j) = 0$ and $cd(g_i, g_j) > 0$, then there exists full interference (--) relation between goals.

Because of the interdependent relations between goals agent should select low-level goals from executable goals to take actions. The selection process is based on goal graph and the relations between different goals. In certain negotiation stage agents can compute the satisfaction degrees of upper level goals by using following formulae (1) and (2).

When there exists AND-relation between goal g_1 and g_2

$$\begin{aligned} & sat_{and}(sat_{g_1}(o), sat_{g_2}(o)) \\ &= \theta_{and} \min\{sat_{g_1}(o), sat_{g_2}(o)\} + (1 - \theta_{and}) \frac{(sat_{g_1}(o) + sat_{g_2}(o))}{2} \end{aligned} \quad (1)$$

where $\theta_{and} = (cd - sd + 1) / 2$

When there exists OR-relation between goals

$$\begin{aligned} & sat_{or}(sat_{g_1}(o), sat_{g_2}(o)) \\ &= \theta_{or} \max\{sat_{g_1}(o), sat_{g_2}(o)\} + (1 - \theta_{or}) \frac{(sat_{g_1}(o) + sat_{g_2}(o))}{2} \end{aligned} \quad (2)$$

where $\theta_{or} = (sd - cd + 1) / 2$

Based on formulas above and goal graph agents can compute the satisfaction degrees of various goals during negotiations. Agents search the goal graph from the root nodes to leaf nodes and select the maximal set of executable goals based on AND/OR-relations between goals. And when low-level goals are selected, agents use negotiation strategies to achieve them. Satisfaction degrees of various goals are computed to control negotiation processes. That is, agents should adjust the set of low-level goals according to negotiation states.

3.3 Goal Execution and Goal Reconsideration

Negotiation strategies direct agent's actions toward goals and provide the negotiation concession modes of agents. Selection of negotiation strategies is influenced by many factors such as goal priorities, relations between goals, negotiation states. Outcomes obtained by using negotiation strategies are uncertain. For example, one agent has made some concessions on one issue and expects other agent can make concessions on the same issue. However, other agent does not make any concession. Certainly, agents can ignore these fluctuations and keep the same negotiation strategies, but agents can achieve a higher satisfaction degree by proactively adjusting goals and strategies. For example, when agents find goal g_i cannot be achieved, they should search goal graph and select another goal g_j that have OR-relation between them.

$CG(\cdot)$ can be used to evaluate strategies. If negotiation strategies are effective, then

the value of $CG(\cdot)$ should increase. Further, when these are interdependent relations between goals, agents can evaluate strategies based on satisfaction degrees of upper level goals by using formulae (1) and (2).

We take the view that agents should reconsider goals when their beliefs or preferences concerning their goals change. Note that agents do not change their beliefs or preferences without reason. That is, their goals will tend to be stable. The more important the goals are, the greater is the motivation to move towards the goals. However, when some goals are impossible to achieve, agents should reconsider their goals and reset them. Goal reconsideration is a complex process and the mechanism must be considered carefully. All agents have to make choices about how to reconsider their mental states at each moment in time. The relations between goals and strategies are controlled by agents' deliberation processes which can be considered as consisting of reasoning about mental states, selecting goals, planning strategies, selecting and executing strategies [13]. The deliberation processes should control whether a goal still exist during the execution of strategies to achieve that goal. If the corresponding goal is reached or dropped with some reasons, then the deliberation processes can change or allow continuing with the strategy. In addition, agents may execute multiple strategies interleaved or consecutively. Interleaving might be beneficial, but mechanisms are needed to solve conflicts between strategies.

4 Conclusions

In this paper we propose a framework for goal driven negotiation process. When users delegate negotiation tasks to agents, agents determine appropriate goals pursued based on relevant information. During negotiations according to intensity values of motivations, agents can activate high-level goals and set goal priorities. High-level goals are decomposed into executable goals. Agents select some executable goals to achieve by using negotiation strategies and reconsider goals when necessary. The framework can be used to design negotiating agents in electronic commerce. However, the framework is designed in the rough and many problems should be solved in the future such as detailed goal selection and reconsideration mechanism, reasoning process of agents.

Acknowledgments. We thank the support of Humanities and Social Science Research Foundation of Ministry of Education of China (No. 07JA630052), and Doctor Foundation of Henan Polytechnic University of China (No. B2009-24).

References

1. Maes, P., Guttman, R.H., Moukas, A.G.: Agents That Buy and Sell. *Communication of the ACM* 42(3), 81–91 (1999)
2. Wurman, P.R., Wellman, M.P., Walsh, W.E.: The Michigan Internet AuctionBot: a configurable auction server for human and software agents. In: *Proceedings of the Second International Conference on Autonomous Agent*, Minneapolis, Minnesota, US, pp. 301–308 (1998)

3. Kraus, S.: Negotiation and cooperation in multi-agent environments. *Artificial Intelligence* 94, 79–97 (1997)
4. Bichler, M., Segev, A.: Methodologies for the design of negotiation protocols on E-markets. *Computer Networks* 37, 137–152 (2001)
5. Rosenschein, J.S., Zlotkin, G.: *Rules of Encounter: Designing Conventions for Automated Negotiation Among Computers*. The MIT Press, Cambridge (1994)
6. Fatima, S.S., Wooldridge, M.J., Jennings, N.R.: Optimal Negotiation Strategies for Agents with Incomplete Information. In: Meyer, J.-J.C., Tambe, M. (eds.) *ATAL 2001*. LNCS (LNAI), vol. 2333, pp. 53–68. Springer, Heidelberg (2002)
7. Greenwald, A.: The 2002 trading agent competition: An overview of agent strategies. *AI Magazine* 24(1), 83–91 (2003)
8. Lopes, F., Mamede, N., Novais, A.Q., Coelho, H.: A negotiation model for autonomous computational agents: Formal description and empirical evaluation. *Journal of Intelligent & Fuzzy Systems* 12, 195–212 (2002)
9. Munroe, S.J., Luck, M.: Agent Autonomy Through the $3M$ Motivational Taxonomy. In: Nickles, M., Rovatsos, M., Weiss, G. (eds.) *AUTONOMY 2003*. LNCS (LNAI), vol. 2969, pp. 55–67. Springer, Heidelberg (2004)
10. Kersten, G.E., Szpakowicz, S.: Negotiation in Distributed Artificial Intelligence: Drawing from Human Experience. In: *Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Sciences*, pp. 258–270 (1994)
11. Shen, Z., Gay, R., Miao, Y., Miao, C.: Goal Autonomous Agent Architecture. In: *Proceedings of the 28th Annual International Computer Software and Applications Conference (COMPSAC 2004)*, vol. 02, pp. 45–46 (2004)
12. Li, H., Stanley, Y., Su, W., Lam, H., Huang, Y.: Automated E-business Negotiation: Model. In: *Life Cycle and System Architecture, Computer and Information Science and Engineering University of Florida, Technical Report* (2001)
13. Kakas, A., Moraitis, P.: Argumentative Deliberation for Autonomous Agents. In: *Proceedings of the ECAI 2002 Workshop on Computational Models of Natural Argumentation*, pp. 65–74 (2002)

Inconsistency in Multi-Agent Systems

Du Zhang

Department of Computer Science, California State University,
Sacramento, CA 95819-6021, USA
zhangd@ecs.csus.edu

Abstract. Multi-agent systems are distributed problem solving systems involving multiple collaborating intelligent agents that are capable of interacting with their environments. Toward the goal of developing multi-agent systems of bounded rationality, an important issue is how to manage and handle inconsistent or conflicting knowledge and information an agent may possess or has to reason with. The focus of this paper is on inconsistency in multi-agent systems. We describe the occurrence of inconsistency in the depth of knowledge, define nine different types of inconsistent phenomena, and discuss possible ways to utilize inconsistency as useful heuristics toward developing multi-agent systems of bounded rationality. The main contribution is that we shed some new light on inconsistency in multi-agent systems.

Keywords: multi-agent systems, inconsistency, bounded rationality, heuristics.

1 Introduction

Multi-agent systems (MAS) are distributed problem solving systems involving multiple collaborating intelligent agents. Each agent in an MAS has its knowledge base, its meta-knowledge on how to utilize knowledge in problem solving, and a set of constraints, assumptions, or defaults from its underlying ontology or problem domain. An agent is capable of interacting with its environment through receiving input data and taking actions that will impact the environment. An agent collaborates with other agents through data and knowledge exchange and fusion.

When engaged in problem solving, agents' decision-making process is confined by the following constraints: the knowledge or information they possess, the cognitive limitations they have, and the time limit within which a decision needs to be made [19, 20]. As a result, the agents' decision-making process in practice really consists in a search through a finite number of options, the fewer the better. Agents need to exploit pre-existing structures or regularities in the environment, apply approximate or heuristic approaches to problems, and be content with good enough solutions. This is the essence of the theory of *bounded rationality* by Simon [20], which also underpins human intelligent behaviors.

In a nutshell, an agent's practical decision-making process is not a perfect rational process of finding an optimal solution (a solution that maximizes the expected utility) given the information available from the environment. Agents of bounded rationality exhibit *satisficing*, rather than *optimizing*, behavior: (a) seeking satisfactory solutions

rather than optimal ones; (b) adopting simplified choices; (c) deliberating only long enough; and (d) relying on heuristic approaches rather than rigid rule of optimization.

Toward the goal of developing bounded rational MAS, a whole host of human cognitive skills and capabilities should be brought to bear. One important capability is how to manage and handle inconsistent or conflicting knowledge and information an agent may possess or has to reason with. Inconsistency is an important phenomenon that exists ubiquitously in human behaviors and in various aspects of real life [13]. Inconsistent phenomena manifest themselves in data, information, knowledge, meta-knowledge, and expertise [29]. Inconsistent or conflicting assumptions, beliefs, evidences, or options can serve as important heuristics in the decision-making process of a bounded rational agent [27, 28].

In this paper, our focus is on inconsistency in MAS. Before adopting any strategies for utilizing inconsistency as heuristics toward developing bounded rational MAS, we need to have a good understanding on the types of conflicting phenomena in MAS.

The rest of the paper is organized as follows. Section 2 offers a brief overview on the occurrence of inconsistency in the depth of knowledge. Section 3 defines nine different types of inconsistent phenomena in MAS. Section 4 discusses how to utilize inconsistency as heuristics for MAS of bounded rationality. Finally, Section 5 concludes the paper with remarks on future work.

2 Inconsistency in Depth of Knowledge

Agents make decisions based on the data, information, knowledge and meta-knowledge they possess or have access to. Inconsistency can arise in various circumstances at different levels in the depth of knowledge, from data, information, knowledge, meta-knowledge, to expertise. Table 1 summarizes features in each level in the depth of knowledge.

Table 1. Levels in Knowledge Depth

Knowledge Depth	Features
Expertise	<i>Compartmentalization, epistemic backgrounds</i>
Meta-knowledge	<i>Control strategies, heuristics, learning decisions</i>
Knowledge	<i>Declarative, procedural</i>
Information	<i>Functional dependencies, association</i>
Data	<i>Symbolic, numeric, graphic, categorical, or waveform valuations</i>

In general, there are some shifting trends in terms of structures, representations, semantics, and entity definitions in the knowledge depth hierarchy, as shown in Figure 1.

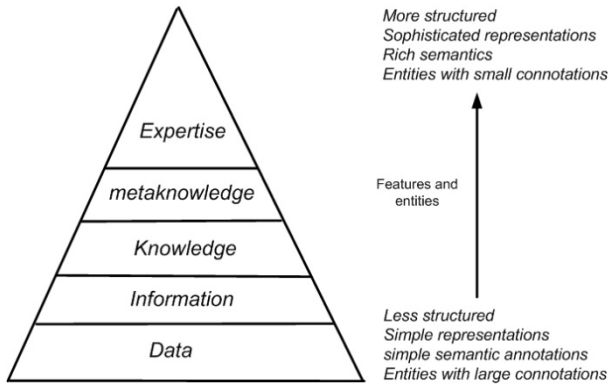


Fig. 1. Shifting Features in Depth of Knowledge

Data inconsistency arises when patterns in data do not conform to an established range, distribution or interpretation. The exponentially growing volumes of data stemming from almost all types of data being created in digital form, a proliferation of sensors and sensor networks, and other sources such as social networks, crowd-sourced web applications, complex computer simulations, space explorations, and high-resolution imagery and video, have all contributed to an ever increasing trend of data inconsistency. At this level, inconsistency is often studied in the context of detecting outliers, anomalies, exceptions, surprises, or peculiarities [7]. The types of data points or entities are numeric, symbolic, graphic, categorical, or waveform.

Information inconsistency occurs when meanings of the same data values become conflicting or when the same attribute for an entity has different data values at different places. At the information level, inconsistency is investigated through scrutinizing functional dependencies and association relationships among entities [10, 14, 23].

Knowledge inconsistency happens when propositions of either declarative or procedural beliefs, in explicit or tacit form, yield antagonistic outcomes for the same set of conditions or circumstances. The proliferation of crowd-sourced knowledge bases in knowledge-rich domains further exacerbates the need for knowledge inconsistency management. At the knowledge level, the focus is on how to understand and recognize various types of conflicting circumstances in knowledge and what theories and approaches can be developed to cope with and circumvent them [24, 25, 26].

Inconsistency can also emerge from meta-knowledge, i.e., knowledge about knowledge (e.g., strategies to reduce overfitting when utilizing certain learning approach to generate an approximation to a target function). Conflicting meta-knowledge can result in domain knowledge being utilized in ways that generate antagonistic outcomes [9]. Very little research on meta-knowledge inconsistency has been done.

Finally, inconsistency can stem from the highest level in the depth of knowledge, expertise, which is specialized operative knowledge that is inherently task-specific and relatively inflexible. Inconsistent expertise can be attributed to epistemic reasons or compartmentalization in the memory apparatus. Much remains to be done to develop some good understanding of what the expertise inconsistency entails.

3 Types of Inconsistency in MAS

To facilitate the discussion, the structure of an MAS is given in Figure 2 where agents work independently or collaborate for a given task. For each agent, it has its knowledge base, its meta-knowledge on how to utilize knowledge in problem solving, and a set of constraints, assumptions, or defaults from its underlying ontology or problem domain. During problem solving sessions, an agent receives input data, and makes a decision at some point in time based on what it possesses and what it acquires through exchanging with other agents.

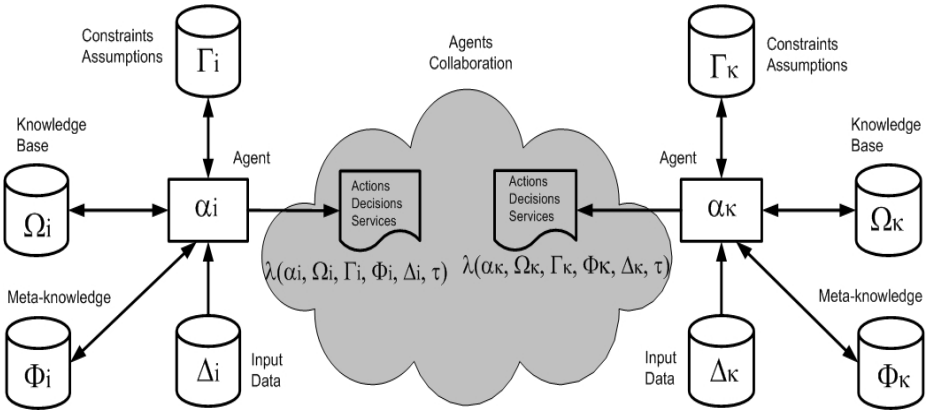


Fig. 2. Multi-Agent System Environment

Definition 1. We adopt the following notations where $i \in [1, \dots, n]$ for $n > 1$:

- α_i : i^{th} agent in an MAS of n agents.
- Ω_i : knowledge base of α_i .
- Γ_i : constraints, assumptions, or defaults from its underlying ontology or problem domain for α_i .
- Φ_i : meta-knowledge of α_i .
- Δ_i : input data to α_i for a given task.
- τ : time instance.
- $\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)$: a decision/an action/a service α_i makes/takes/provides is a function of the agent's knowledge base, domain constraints, meta-knowledge, input data, and time instance.
- $[\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)]$: denotes the process through which α_i arrives at a decision at τ based on its Ω_i , Γ_i , Φ_i , and Δ_i .
- $\mathfrak{S}(\Omega_i), \mathfrak{S}(\Gamma_i), \mathfrak{S}(\Phi_i), \mathfrak{S}(\Delta_i)$: indicates that there exists some inconsistency in Ω_i , Γ_i , Φ_i , or Δ_i , respectively.
- $L_i \models L_k$: indicates that L_i logically implies L_k , where L_i and L_k are literals (or conjunctions of literals).
- $L_i \not\models L_k$: represents that L_i and L_k are inconsistent with each other. □

Definition 2. With regard to the multi-agent system environment in Figure 2 and the aforementioned notations, we define four categories of inconsistency:

- *Data inconsistency*: $\mathfrak{S}(\Delta_i)$.
- *Knowledge inconsistency*: $\mathfrak{S}(\Omega_i) \vee \mathfrak{S}(\Gamma_i)$.
- *Meta-knowledge inconsistency*: $\mathfrak{S}(\Phi_i)$.
- *Decision inconsistency*: $[\lambda(\dots) \Leftrightarrow \lambda(\dots)] \vee [\lambda(\dots) \Leftrightarrow \Delta_i] \vee [\lambda(\dots) \Leftrightarrow \Gamma_i]$. \square

In the remainder of this section, we define nine different types of inconsistency in MAS, focusing on how different inconsistent circumstances manifest themselves in the MAS decision process.

3.1 Reasoning in Presence of Knowledge Inconsistency

An agent α_i makes a decision in the presence of inconsistency in Ω_i (knowledge inconsistency).

$$\mathfrak{S}(\Omega_i) \wedge [\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)]$$

Under this circumstance, α_i is aware of the fact that its Ω_i contains inconsistency, and circumvents the *explosive reasoning* in its decision process by incorporating one of the several possible approaches [3]: (a) dealing with maximal consistent subsets in Ω_i ; (b) adopting multi-valued logic; (c) relying on defeasible logic in its nonmonotonic reasoning [2]; (d) weighing on arguments and counter-arguments, and their strengths [4, 8].

3.2 Inconsistent Decisions

Consecutive decisions an agent α_i made at τ and τ' are inconsistent with each other (decision inconsistency).

$$[\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)] \wedge [\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau')] \\ \wedge [\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau) \Leftrightarrow \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau')] \wedge [\tau \neq \tau']$$

Such a decision behavior in α_i can be useful and effective in *inconsistent deception*, one of the defensive techniques to fend off attackers [16]. Central to the approach is to generate conflicting responses to an attacker's requests in order to confuse or disorient the attacker.

3.3 Decisions Inconsistent with Δ_i

An agent α_i 's decision is inconsistent with Δ_i (decision inconsistency).

$$[\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)] \wedge [\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau) \Leftrightarrow \Delta_i]$$

A recent success story on utilizing inconsistency as an effective tool in *open domain question-answering* agent systems is the IBM's Watson system [12]. Watson was developed with the goal to compete at the human champion level in real time on *Jeopardy!*, an American TV quiz show that dates back in 1964. Early in 2011, Watson competed in the show's first ever human-versus-machine matchup and was pitted

against two top human players. Watson was triumphant over the human players and walked away with the first prize of one million dollars.

A cornerstone to Watson's open-domain question-answering capability is IBM's *DeepQA* technology that allows for question analysis and decomposition, hypothesis generation and filtering, evidence retrieval and scoring, merging and ranking answers, and confidence estimation [12]. Several components in Watson use inconsistency as heuristics for hypothesis and evidence scoring. For example, temporal reasoning is deployed in Watson to "detect inconsistencies between dates in the clue and those associated with a candidate answer" [12]. Another scoring component, geospatial reasoning, is capable of detecting spatial inconsistencies stemming from conflicting spatial relations such as "directionality, borders, and containment between geonities" [12]. During the process to combine scores for individual pieces of evidence to produce overall evidence profiles for candidate answers, Watson examines inconsistencies from dimensions such as name consistency or theory consistency [12].

Despite the aforementioned inconsistency-based heuristics in place in its DeepQA engine, Watson flubbed during the Final Jeopardy! question at the end of the first match. The category for the Final Jeopardy! question was "US Cities," and the answer was: "Its largest airport is named for a World War II hero; its second largest, for a World War II battle." The given category and the answer for the Final Jeopardy! question constitutes the input (Δ_i) to Watson (α_i). Whereas both human players responded correctly with "What is Chicago?", Watson's response was a confused one: "What is Toronto?????".

The case of Watson's response of "Toronto" to the question category of "US Cities" constitutes what is referred to as *anti-subsumption* inconsistency in [24, 26] (given a taxonomy of concepts, when there is no longer a subsumed relationship between a sub-concept, Toronto, Canada being a US city with airports, and its super-concept, US Cities). This antagonistic circumstance is a salient case of decision inconsistency with Δ_i .

3.4 Decisions Inconsistent with Γ_i

An agent α_i 's decision is inconsistent with Γ_i (decision inconsistency).

$$[\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)] \wedge [\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau) \not\Leftarrow \Gamma_i]$$

In default reasoning, an agent may need to augment its set of base beliefs in Ω_i with the negation of any ground atom that is not entailed by Ω_i . This is what is referred to as the *closed-world assumption* (CWA) which can be part of the agent's Γ_i . CWA allows for an augmented theory about a domain that consists of the original set of base beliefs plus a set of negative ground literals.

An agent's knowledge is *complete* if and only if for every sentence β in its vocabulary, either β or $\neg\beta$ is known [6]. In general, an agent's knowledge can be incomplete. For instance, if an agent's knowledge contains the following: $P(a) \vee Q(a)$, then the agent cannot deduce either $P(a)$ or $\neg P(a)$ (neither can it deduce $Q(a)$ or $\neg Q(a)$) without CWA. However, with CWA in its Γ_i , α_i would produce decisions or actions that can be inconsistent. For example, given the sentences $\{P(x) \vee Q(x), P(a), Q(b)\}$ in Ω_i ,

if a problem domain contains three constants $\{a, b, c\}$, then α_i would generate the following under CWA: $\{P(c) \vee Q(c), \neg P(c), \neg Q(c)\}$, which is inconsistent.

3.5 Data Inconsistency

An agent α_i 's decision indicates that there is inconsistency in Δ_i (data inconsistency).

$$[\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)] \wedge [\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau) \models \mathfrak{I}(\Delta_i)]$$

This type of inconsistency can be better illustrated through some real-world applications such as digital image forensics. Given digital images as Δ_i , there are various inconsistency-based heuristics an agent can utilize to detect image forgery or tampering.

The first type of data inconsistency in digital image forensics is referred to as *lighting inconsistency* [11]. In digital images, lighting inconsistency highlights a circumstance where shapes, colors, locations or directions, or number of sources of reflected light on objects are inconsistent. For instance, when specular highlights (small white specks of reflected light in people's eyes) in an image result in inconsistent lighting directions, number of sources, or shapes, it is evidence of image altering [11]. The second type of data inconsistency in digital image forensics is the *local noise level inconsistency* [15]. Normally in an authentic digital image, the noise is uniformly distributed across the entire image. Tools used to alter images often inject locally randomized noise to the forged regions so as to conceal the evidence of tampering, thus creating inconsistencies in the levels of noise in the image. The third type of data inconsistency in digital image forensics is the *blocking artifact inconsistencies* [22]. In JPEG images, due to the fact that digital camera manufacturer and image processing software often utilize different JPEG quantization tables to balance between compression ratio and image quality, different blocking artifacts will be injected into the images as a result. When a digitally forged image is created from several sources, the resulting image invariably contains different sorts of compression artifacts. In addition, many image manipulation operations such as image splicing, re-sampling, or skin optimization, will generate differential blocking artifacts. There are additional inconsistency based techniques for detecting tampered images [17, 21].

3.6 Knowledge Inconsistency Due to Services

An agent α_i 's decision indicates that there is inconsistency in Ω_i (knowledge inconsistency).

$$[\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)] \wedge [\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau) \models \mathfrak{I}(\Omega_i)]$$

There are circumstances where the services an agent provides are at the expense of knowledge inconsistency. For instance, in a Unix environment, a non-root user cannot directly access the password file, but can nevertheless invoke a special *setuid* program that temporarily elevates the user's privilege for the task of changing password (thus updating the system's password file). Normally a user's effective uid (**euid**) is the same as the real uid (**ruid**) which is a non-zero value assigned by the system (**ruid** = **euid**). As the user is executing this *setuid* program for password changing purpose, the **euid** becomes zero (root privilege so that the password file can be changed as a result of this user changing her password) while the **ruid** remains as the initially

assigned non-zero value, therefore creating what is referred to as the *setuid inconsistency* (when a process' *ruid* \neq *euid*) [30]. This identity inconsistency is regarded as an inconsistency in Ω_i as a result of the system agent providing a service to users. It is transient as the *euid* will be set to that of the *ruid* once the service is done.

3.7 Inconsistent Constraints

An agent α_i 's decision indicates that there is inconsistency in Γ_i (knowledge inconsistency).

$$[\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)] \wedge [\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau) \models \mathfrak{S}(\Gamma_i)]$$

The fact that there can be inconsistency in Γ_i stems from the possibilities that models and constraints, assumptions, or defaults behind models in Γ_i often times may be incompatible, conflicting or inconsistent with each other. For instance, if we have two assumptions in Γ_i regarding the precedence relationship between software requirements and implementation: one is based on the *Waterfall* process model and states that “*requirements are completely specified before implementation*” (γ_{wf}) [1], and the other is based on the success model of *IKIWISI* (I'll know it when I see it) [5] and states that “*implementation exists for some requirements*” (γ_{iki}). Capturing these two assumptions in the first-order logic, we have the following [25]:

- γ_{wf} : which can be translated into the following:
 $\forall x \forall y [Requirement(x) \wedge Implementation(y) \supset Precedence(x, y)]$
- γ_{iki} : which can be translated into:
 $\exists x \exists y [Requirement(x) \wedge Implementation(y) \wedge Precedence(y, x)]$

γ_{wf} and γ_{iki} can be transformed into the following clauses where A , and B are ground terms from the domain that denote a specific piece of requirement and a specific implementation, respectively:

$$\begin{aligned} &\neg Requirement(x) \vee \neg Implementation(y) \vee \neg Precedence(y, x) \\ &Requirement(A) \\ &Implementation(B) \\ &Precedence(B, A) \end{aligned}$$

The derivation of an empty clause in the resolution tree (Figure 3) indicates that the two are inconsistent.

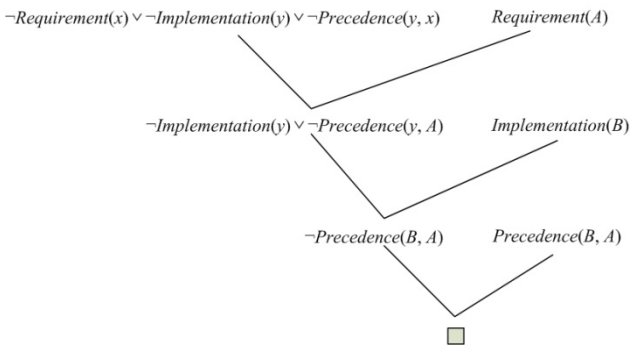


Fig. 3. Resolution tree for γ_{wf} and γ_{iki} .

3.8 Decision Inconsistency by Multiple Agents

Decisions agents α_i and α_k make during their collaboration are inconsistent (decision inconsistency).

$$\begin{aligned} & [\alpha_i \vdash \lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau)] \wedge [\alpha_k \vdash \lambda(\alpha_k, \Omega_k, \Gamma_k, \Phi_i, \Delta_k, \tau')] \\ & \wedge [\lambda(\alpha_i, \Omega_i, \Gamma_i, \Phi_i, \Delta_i, \tau) \Leftrightarrow \lambda(\alpha_k, \Omega_k, \Gamma_k, \Phi_i, \Delta_k, \tau')] \wedge [\alpha_i \neq \alpha_k] \end{aligned}$$

In a multi-agent system, autonomous agents have their own local knowledge and meta-knowledge, and their own constraints, assumptions, or defaults about the worlds they are in. when agents collaborate through exchanging data or information, or fusing knowledge, to arrive at a decision, inconsistency can arise as a result of their *epistemic* differences.

3.9 Meta-Knowledge Inconsistency

Let $\Phi_i = \{h_i \mid h_i \text{ is a heuristic on how } \Omega_i \text{ or } \Gamma_i \text{ is utilized}\}$. We can regard $h_{im} \in \Phi_i$ as a mapping from $\langle \Omega_i, \Gamma_i \rangle$ to $\langle \Omega_{im}, \Gamma_{im} \rangle$, where $\Omega_{im} \subseteq \Omega_i$ and $\Gamma_{im} \subseteq \Gamma_i$:

$$h_{im} : \langle \Omega_i, \Gamma_i \rangle \rightarrow \langle \Omega_{im}, \Gamma_{im} \rangle.$$

Given $h_{ij} \in \Phi_i$ and $h_{ik} \in \Phi_i$, and $h_{ij} \neq h_{ik}$, we have

$$\begin{aligned} h_{ij} & : \langle \Omega_i, \Gamma_i \rangle \rightarrow \langle \Omega_{ij}, \Gamma_{ij} \rangle, \text{ where } \Omega_{ij} \subseteq \Omega_i \text{ and } \Gamma_{ij} \subseteq \Gamma_i; \text{ and} \\ h_{ik} & : \langle \Omega_i, \Gamma_i \rangle \rightarrow \langle \Omega_{ik}, \Gamma_{ik} \rangle, \text{ where } \Omega_{ik} \subseteq \Omega_i \text{ and } \Gamma_{ik} \subseteq \Gamma_i. \end{aligned}$$

If two decisions an agent α_i makes are guided by h_{ij} and h_{ik} respectively for a task, and the resulting decisions are conflicting, then this is attributed to meta-knowledge (meta-knowledge inconsistency).

$$\begin{aligned} & [\alpha_i \vdash \lambda(\alpha_i, \Omega_{ij}, \Gamma_{ij}, h_{ij}, \Delta_i, \tau)] \wedge [\alpha_i \vdash \lambda(\alpha_i, \Omega_{ik}, \Gamma_{ik}, h_{ik}, \Delta_i, \tau)] \\ & \wedge [\lambda(\alpha_i, \Omega_{ij}, \Gamma_{ij}, h_{ij}, \Delta_i, \tau) \Leftrightarrow \lambda(\alpha_i, \Omega_{ik}, \Gamma_{ik}, h_{ik}, \Delta_i, \tau)] \end{aligned}$$

Assuming we have the following Ω_i for classifying animals and plants:

- R_1 : $SeaCucumber(x) \rightarrow CucumberShape(x)$
- R_2 : $CucumberShape(x) \rightarrow Cucumber(x)$
- R_3 : $Cucumber(x) \rightarrow Plant(x)$
- R_4 : $SeaCucumber(x) \rightarrow NonRigidCellWalls(x)$
- R_5 : $NonRigidCellWalls(x) \rightarrow Animal(x)$

Let $\Phi_i = \{h_1, h_2\}$ where h_1 indicates a morphology based approach for classification (hence involving R_1 , R_2 , and R_3), and h_2 prescribes a cell structure based approach for classification (thus involving R_4 , and R_5). If we have $\Delta_i = \{SeaCucumber(D)\}$, then h_1 yields a classification of $Plant(D)$ and h_2 leads to a classification of $Animal(D)$. $Plant(D)$ and $Animal(D)$ represent a *mutual exclusive inconsistency* [24, 26].

4 Inconsistency as Heuristics for MAS of Bounded Rationality

Recognizing the presence of various types of inconsistency during agents' decision making process is just the first step. The objective is how to utilize inconsistency as useful heuristics in helping achieve an agent's bounded rationality in its problem solving process.

For agents to reason in the presence of inconsistency, several approaches are viable: (a) *paraconsistent logic* that allows agents to either deal with maximal consistent subsets in Ω_i or adopt a multi-valued logic approach when reasoning with inconsistent knowledge; (b) *defeasible logic* and its *nonmonotonic reasoning*; and (c) *argumentation* based approach that weighs on the strengths of arguments and counter-arguments.

For agents' inconsistent decisions, some can be desirable as discussed in subsection 3.2, while others demonstrate the *cognitive penetrability* [18]. Meta-knowledge can be learned to differentiate desirable inconsistent decisions from the undesirable ones. This can be important in supporting the objectives of bounded rationality (exploiting pre-existing structures or regularities in the environment, and applying approximate or heuristic approaches to problems).

For an agent's decisions inconsistent with data or inconsistent decisions by multiple agents, they can be utilized as an impetus for the agents involved to be engaged in the inconsistency-induced learning [28].

Data inconsistency can serve as effective heuristics in problem solving process. It provides direct support to the objectives of bounded rationality.

Meta-knowledge has impact on the satisficing behaviors of seeking satisfactory solutions rather than optimal ones, adopting simplified choices, and deliberating only long enough. Meta-knowledge inconsistency also affords an agent the opportunity to refine its meta-knowledge through the inconsistency-induced learning [28].

5 Conclusion

In this paper, we highlight the issue of how to delineate different types of inconsistent circumstances in MAS so as to take full advantage of utilizing inconsistency as meaningful and effective heuristics in helping accomplish the bounded rationality for agents in their problem solving processes. The main contributions of this work lie in the following: (a) we adopt a holistic view toward inconsistency in MAS (not every inconsistency is bad, some are even desirable); (b) we define nine different types of inconsistent phenomena for MAS; (c) we establish the connection between utilizing inconsistency as useful heuristics and the bounded rationality.

Future work can be pursued in the following directions. Further development of inconsistency based heuristics is needed. Augmenting the existing MAS structure to include some inconsistency-induced learning component is very appealing. Finally empirical studies need to be conducted to substantiate the utility and viability of the proposed approach.

References

1. Al-Said, M.: Identifying, Analyzing, and Avoiding Software Model Clashes, Ph.D. Dissertation. University of Southern California (2003)
2. Antoniou, G., Billington, D., Governatori, G., Maher, M.: Representation results for de-feasible logic. *ACM Trans. On Computational Logic* 2(2), 255–286 (2001)
3. Bertossi, L., Hunter, A., Schaub, T.: Introduction to inconsistency tolerance. In: Bertossi, L., Hunter, A., Schaub, T. (eds.) *Inconsistency Tolerance*. LNCS, vol. 3300, pp. 1–14. Springer, Heidelberg (2005)
4. Besnard, P., Hunter, A.: *Elements of Argumentation*. MIT Press (2008)
5. Boehm, B.: Requirements that handle IKIWISI, COTS, and rapid change. *IEEE Computer* 33(7), 99–102 (2000)
6. Brachman, R.J., Levesque, H.J.: *Knowledge Representation and Reasoning*. Morgan Kaufmann, San Francisco (2004)
7. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. *ACM Computing Surveys* 41(3) (2009)
8. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77, 321–357 (1995)
9. Evans, J.A., Foster, J.G.: Metaknowledge. *Science* 331, 721–725 (2011)
10. Fan, W., Geerts, F., Lakshmanan, L., Xiong, M.: Discovering conditional functional dependencies. In: *Proc. of the International Conference on Data Engineering (ICDE 2009)*, pp. 1231–1234 (2009)
11. Farid, H.: Seeing is not believing. *IEEE Spectrum* 46(8), 44–51 (2009)
12. Ferrucci, D., et al.: Building Watson: an overview of the DeepQA project. *AI Magazine*, 59–79 (Fall 2010)
13. Gotesky, R.: The uses of inconsistency. *Philosophy and Phenomenological Research* 28(4), 471–500 (1968)
14. Martinez, M.V., Pugliese, A., Simari, G.I., Subrahmanian, V.S., Prade, H.: How Dirty Is Your Relational Database? An Axiomatic Approach. In: Mellouli, K. (ed.) *ECSQARU 2007*. LNCS (LNAI), vol. 4724, pp. 103–114. Springer, Heidelberg (2007)
15. Mahdian, B., Saic, S.: Using noise inconsistencies for blind image forensics. *Image and Vision Computing* 27, 1497–1503 (2009)
16. Neagoe, V., Bishop, M.: Inconsistency in deception for defense. In: *Proc. of the 2006 Workshop on New Security Paradigms*, Schloss Dagstuhl, German, September 19–22, pp. 31–38 (2006)
17. Popescu, A.C.: *Statistical Tools for Digital Image Forensics*, Ph.D. thesis, Department of Computer Science, Dartmouth College (December 2004)
18. Pylyshyn, Z.: Computing in cognitive science. In: Posner, M.I. (ed.) *Foundations of Cognitive Science*, pp. 49–92. MIT Press, Cambridge, MA (1989)
19. Russell, S., Norvig, P.: *Artificial Intelligence: A Modern Approach*. Prentice-Hall (2010)
20. Simon, H.A.: *Models of Bounded Rationality*. MIT Press (1982)
21. Wang, X., Xuan, B., Peng, S.: Digital image forgery detection based on the consistency of defocus blur. In: *Proc. of International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 192–195 (2008)
22. Ye, S., Sun, Q., Chang, E.C.: Detecting digital image forgeries by measuring inconsistencies of blocking artifacts. In: *Proc. of the IEEE International Conference on Multimedia and Expo*, pp. 12–15 (2007)

23. Yeh, P., Puri, C.: An efficient and robust approach for discovering data quality rules. In: Proc. of the 22nd IEEE International Conference on Tools with Artificial Intelligence, Arras, France, pp. 248–255 (2010)
24. Zhang, D.: Taming Inconsistency in value-based software development. In: Proceedings of 21st International Conference on Software Engineering and Knowledge Engineering, Boston, Mass., pp. 450–455 (July 2009)
25. Zhang, D.: Capturing antagonistic stakeholders value propositions in value-based software development. In: Proc. of 22nd International Conference on Software Engineering and Knowledge Engineering, San Francisco, pp. 12–17 (July 2010)
26. Zhang, D.: Toward a classification of antagonistic manifestations of knowledge. In: Proc. of the 22nd IEEE International Conference on Tools with Artificial Intelligence, Arras, France, pp. 375–382 (October 2010)
27. Zhang, D.: Inconsistency-induced heuristics for problem solving. In: Proc. of the 23rd International Conference on Software Engineering and Knowledge Engineering, Miami, FL, pp. 137–142 (July 2011)
28. Zhang, D., Lu, M.: Inconsistency-induced learning: a step toward perpetual learners. In: Proc. of 10th IEEE International Conference on Cognitive Informatics, Banff, Canada, pp. 59–66 (August 2011)
29. Zhang, D., Grégoire, É.: The landscape of inconsistency: a perspective. To appear in *International Journal of Semantic Computing*
30. Zhang, D.: The utility of inconsistencies in information security and digital forensics. In: Özyer, T., et al. (eds.) *Recent Trends in Information Reuse and Integration*, pp. 381–397. Springer, Heidelberg (2011)

Part II

**Pattern Recognition, Image
and Video Processing**

Emotional Speech Recognition Based on Syllable Distribution Feature Extraction

Haiying Zhang

Software Engineering Department, Software School of Xiamen University,
Haiyun, Xiamen, 361005, China
zhang2002@xmu.edu.cn

Abstract. With a very broad application prospect, emotional speech recognition has aroused many researchers attentions. The key to the problem is how to extract the features which can express feelings exactly. In this paper, the extraction method based on speech syllable distribution is presented. In contrast with the traditional one the new method can reduce the statistical error without computing the speech pause, moreover, the features such as syllable time and mute time are introduced to aiding emotional speech recognition.

Keywords: emotional speech recognition, feature extraction, NN.

1 Introduction

As the main communication means speech signal is the essential approach to convey information. The signal not only presents the actual semantics but the speakers abundant feelings exploiting rising and falling in cadence. Consequently, it is the main content of emotion computation to distinguish the speakers' emotion according to speech signal. It has a broad prospect such as the human-computer interaction [1], the analysis of customer's degree of satisfaction[2], lie-detect aided[3] etc. however, it is much difficult to extract the features which can distinguish different emotions of speech. In the traditional feature extraction method rhythm and tune are always introduced to assist speech recognition. However, it will bring about statistical error without considering the breaks between syllable. In this paper a new method based on speech syllable distribution characteristic is given and it can be shown by experiments the new one is more precise than the traditional one.

2 Algorithm

2.1 Traditional Feature Extraction Method Based on Rhythm

The research of emotional speech recognition is centered on rhythm and tune [4]. The rhythm is regarded as the most important speech feeling feature such as fundamental ,energy, rate etc. however, speech signal is non-stationary signal, the rhythm has great variation with time. But during short time interval it can be treated as constant relatively. In the traditional method the sentence is firstly decomposed into isometric continuous, short-time segments [5] and then the speech features such as fundamental

frequency, energy are calculated. Finally the statistic vector of the feature is created (for example: max, min, mean, var, etc).

The method has the following demerits:

1. A sentence is composed of several syllable which has break between of them. The break also is involved to computation and it will have great impact on the precision, such as the sadness sentence.
2. The calculation of the break has no statistical significance. During the silence time the amplitude is zero leading to the period tending to zero, so the frequency is ∞ . It is nonsense for emotional speech recognition.

2.2 New Method Presented in This Paper

To overcome the disadvantages of the traditional one, in this paper the feature extraction based on speech syllable distribution characteristic is presented. The main idea is as follows:

- dividing a paragraph of emotion speech into frames;
- distinguish the syllable and non-syllable according to energy;
- analyze the syllable distribution characteristic according to the energy based on frame;
- calculating the rhythm characteristic of the syllable and its statistic parameters.

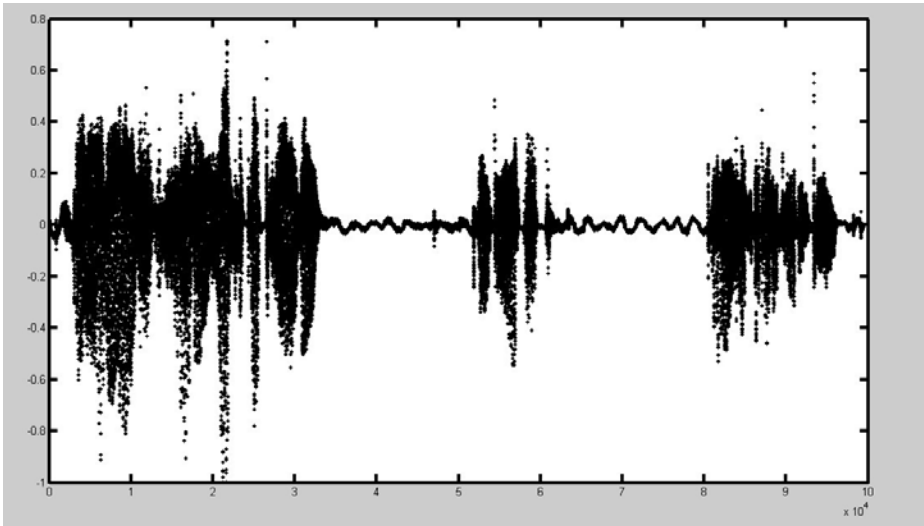


Fig. 1. Syllable distribution characteristic of sad ness speech signal segment

It can be found that different emotion differs not only in frequency, energy also in syllable length and pause time. Such as the syllable of disgust speech segment is shorter than the calmness one, the pause time in sadness speech signal segment is longer than the others. (See in Fig.1).

Consequently, besides the fundamental frequency, energy are involved during feature extraction based on the syllable distribution characteristic ,the syllable mean length, max length, max pause time and mean pause time are also involved. The precision of feature extraction method is key to emotional speech recognition. The speech feature vector is 13D defined in this paper which includes the fundamental frequency and energy statistics such as: mean, maximum, minimum, variance and speed of speech, the length of longest syllable, mean length of syllable, longest pause time, mean pause time etc.

2.3 Syllable Discrimination Algorithm Base on Energy Profile

2.3.1 Main Idea

During the feature extraction based on speech syllable the most important is how to discriminate the syllable one and the non-syllable one in a speech signal segment. We can see from Fig.2.(a)(b), the short-time energy of speech signal can reflect the distribution characteristic of speech syllable. The main idea is as follows:

- Calculate the short-time energy of different frames;
- Compare the energy of adjacent frames one by one ;
- Determine the beginning of a syllable if the energy of continuous n frames are increasing progressively;
- After the beginning of syllable is defined searching one frame backward whose energy is times of the peak energy (threshold is take 0.2) and the frame is away from the beginning one don't exceed the fixed min frame. The frame is the end of the syllable.

2.3.2 Algorithm Description

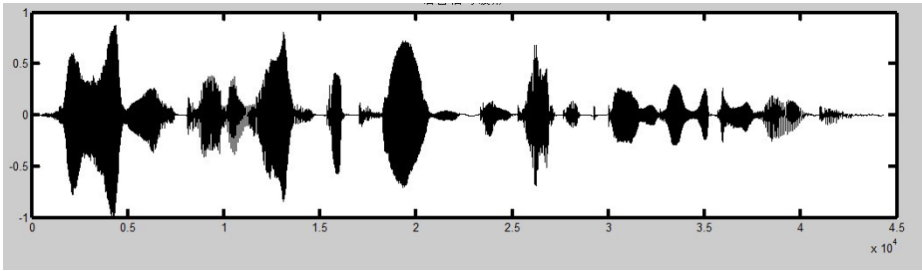
The key of the method is to obtain the frame with max energy and then searching for the frame whose energy is *threshold* time *s* of the max one. Besides this in order to suppress noise the definition of *minVoiceFrameCount* is given to distinguish the noisy syllable. For the speech syllable of human will hold a short time maybe hundred of *ms* but the noise may be 10ms. For the syllable with short sustained time according to the algorithm it will be regarded as noise to remove. after the algorithm is executed the distribution of the syllable is kept in array marked by *Speech Status*. The length of the array means the frames after decomposition.

2.3.3 Design of Emotional Speech Classifier Based on BP NN

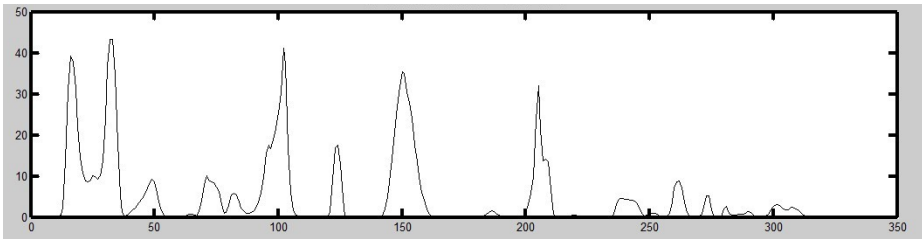
The classifier is designed by BP NN with traditional 3-layer neuron structure, namely input-layer, hidden and output-layer. Four kinds of commonly used emotional speech segments are adopted in this paper such as joy, sadness, disgust and calmness. The neuron in input layer is 13 which means speech feature extracted as above-mentioned algorithm. The neuron in hidden layer is achieved by formula (1):

$$n_{hidden} = \sqrt{n_{input} + n_{output}} + \alpha (1 \leq \alpha \leq 10) \quad (1)$$

The four neuron in output layer means four feelings in training set. If the training feeling is disgust the output vector is $[1 \ 0 \ 0 \ 0]^T$, if it is joy, the output is $[0 \ 1 \ 0 \ 0]^T$ and so on.



(a) wave of a speech signal segment



(b) energy profile of (a)

Fig. 2. Profile of a speech signal segment and its energy distribution

3 Experiments and Analysis

3.1 Emotional Speech DB

By comparison the EMODB from acoustics department in The Technical University of Berlin is defined to be the sample DB.

The emotional speech in EMODB is recorded by 10 professional actors (five actors and five actresses) in a quiet recording studio. The recording material is 10 daily German, including five long sentences and five short sentences. When the EMODB is constructed it simulates seven kinds of human feelings: angry, boredom, disgust, fear, joy, sadness, neutral etc. It includes 800 recordings.

To insure the quality and naturalness of the recording, the emotional speech recording with 80% discrimination ratio and over 60% naturalness are kept to carry out analysis further more. After choice 535 recordings are kept in DB.

In this paper four kinds of emotional speech (disgust, joy, sadness, and calmness) are chosen to be testing samples from 535 speech segments. It is 339 altogether. They are divided into two parts, training set including 200 speech segments and testing set including 139 speech segments.

3.2 Results and Analysis

The classifier is designed by BP NN described in chapter 2. The comparison results using the traditional and the method given in this paper is shown in Tab.1. It can be

shown that the new method presented in this paper has much higher recognition ratio compared with the traditional one in view of the four emotional speech samples. The total ratio is increased from 71.9% to 79.9%. In hence, the new method is superior to the traditional one.

Table 1. Recognition ratio comparison of two kinds of feature extraction methods

		Output			
		disgust	joy	sadness	calmness
input					
	disgust	Traditional	40	11	0
New method		44	8	0	0
joy	Traditional	13	17	0	2
	New method	9	21	0	2
sadness	Traditional	2	2	18	4
	New method	2	1	20	3
calmness	Traditional	0	0	4	25
	New method	0	0	3	26

Acknowledgements. The paper is supported by the Fundamental Research Funds for the Central Universities (0680-ZK1006).

References

1. Lin, L.Y., Gang, W., Cai, Y.K.: The research progress of the recognition of emotional speech. *Journal of Circuits and Systems* 12, 90–98 (2007)
2. Batliner, A., Fischer, K., Huber, R., et al.: How to Find Trouble in Communication. *Speech Communication* 40(1-2), 117–143 (2003)
3. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., et al.: Emotion Recognition in Human-Computer Interaction. *IEEE Signal Processing Magazine* 18(1), 32–80 (2001)
4. Gobl, C., Chasaide, A.N.: Testing Affective Correlates of Voice Quality through Analysis and Resynthesis. In: *ISCA Workshop on Speech & Emotion*, Northern Ireland, pp. 178–183 (2000)
5. Qun, Z.X.: *Decoding of digital speech signal*. China Machine Press (2007)
6. Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., Weiss, B.: *A Database of German Emotional Speech*. *Inter speech* (2005)

Face Recognition Based on the Second-Generation Curvelet Transform Domain and KPCA

Xian Wang, Xin Mu, Yan Zhang, and Fangsheng Zhang

Key Laboratory of Advanced Process Control for Light Industry Ministry
of Education (Jiangnan University), Wuxi, 214122, China
wwxx.2008@163.com, muxin505@sina.com

Abstract. Since wavelet transform can not fully describe facial curves features, in this paper, we propose a novel face recognition method based on Curvelet domain and kernel principal component analysis (KPCA). Using multi-scale, multi-directional Curvelet transform to extract image features not only has higher approximation accuracy and better performance of sparse expression, but also can effectively express the singularity along the curve. Furthermore, kernel principal component analysis (KPCA) is used to project Curvelet feature coefficient into kernel space with more expressing capability. Finally, the nearest method is adopted to classify. The results indicate that the algorithm is effective in image dimension reduction and face recognition rate in the JAFFE face database, ORL face database and FERET face database.

Keywords: face recognition, Curvelet transform, kernel principal component analysis (KPCA), kernel space.

1 Introduction

Over the past 20 years, face recognition technology has been rapidly developed. After Emmanuel and Donoho intuitively introduced theories of the Ridgelet transform, single scale Ridgelet transform and Curvelet transform, the multi-scale analysis in image processing has been widely applied[1-4]. Since the traditional wavelet transformation can only express the singularity of the point in the image, which impacts on wavelet coefficients, the wavelet is difficult to achieve satisfying results of curve expression[5-7]. The main idea of Curvelet transform is, when segmenting the curve, to treat each short segment of curve as a line. Furthermore, the curve information of the image can be expressed by Curvelet transform in multi-scale and multi-direction which allows for obtaining the curve coefficients for effective features. There are various conventional analysis methods used for face image identification through the transformation in the time and frequency domain and combination with the subspace methods, such as the PCA(Principal Components Analysis) and LDA(Linear Discriminant Analysis) [8-10].

Wu Xian-xing [13] employed the Curvelet transform with PCA into the face database for the first time where the Curvelet coefficients were extracted by the PCA approach. However, the conventional PCA tends to be categorized into linear transformation where certain nonlinear information may be missed when the data involves the

characteristics of nonlinear. In order to overcome this obstacle and further analyze of Curvelet coefficient, this work tries to propose a method to process the Curvelet coefficients combined with the kernel principal component analysis (KPCA)[11-12] which may provide more image features. Also, the algorithm employs the kernel function polynomial and obtains the optimal kernel parameter d through experimental tests in the ORL, JAFFE, and FERET face database. Compared with certain conventional methods like PCA with Curvelet, the proposed method shows the better overall performance.

2 Face Information Process with Curvelet Transform

Curvelet transform is actually derived from the ridge wave theory; which consists of a specific filtering process and multiscale Rridgelet transform. A series of filters: Φ_0, Ψ_{2^s} ($s=0, 1, 2, \dots$) are utilized which have to satisfy the following condition:

- 1) Φ_0 is a low-pass filter, whose range of the bandpass is $|\xi| \leq 2$.
- 2) Ψ_{2^s} is a band-pass filter, whose range of the bandpass is $|\xi| \in [2^{2s-1}, 2^{2s+3}]$.
- 3) All filters should satisfy the equation: $|\hat{\Phi}_0(\xi)|^2 + \sum_{s \geq 0} |\hat{\Psi}_{2^s}(\xi)|^2 = 1$.

The essentially inherent relation in Curvelet transform can be demonstrated as *width* \propto *length*² called anisotropic scale relation which is highly directional [1].

Two different methods have been investigated to achieve second-generation digital Curvelet transform before. One is unequally Spaced Fast Fourier Transform (USFFT), while the other one is wrapping algorithm. In comparison with the first generation Curvelet transform, these new discrete Curvelet transform are simpler, faster with less redundancy. In this paper, we use wrapping method. Specific implementation steps are as follows [12]:

1) Apply the 2D FFT (Fast Fourier Transform) where Fourier samples $\hat{f}[n_1, n_2]$ are obtained;

2) multiply the discrete localizing $\hat{U}_{j,l}[n_1, n_2]$ by parameters of scale j and angle l ;

3) Wrap the product above around the origin and obtain

$$\hat{f}_{j,l}[n_1, n_2] = W(\hat{U}_{j,l} \hat{f})[n_1, n_2], \text{ where } 0 \leq n_1 \leq L_{1,j}, 0 \leq n_2 \leq L_{2,j}; L_{1,j} \sim 2^j, \\ L_{2,j} \sim 2^{2j}.$$

4) Apply the inverse 2D FFT to $\hat{f}_{j,l}[n_1, n_2]$, hence the discrete coefficients $C^d(j,l,k)$ is derived.

It is noted that we take the coarse layer coefficient as facial features since the coarse layer of Curvelet coefficients (85*85) can express profile information of human face precisely which is shown in Fig.1.



Fig. 1. The first layer of Curvelet coefficients to recover face images and Original face images

3 Kernel Principal Component Analysis (KPCA)

Rahman S [8] employed conventional PCA method to extract features from Curvelet coefficients. However, the nonlinear part of initial data is missing after using PCA. To overcome the obstacle, KPCA is introduced whose main idea is: extending the linear principal component analysis to nonlinear principal component analysis. The procedure can be briefed here: first samples are mapped into a high-dimensional feature space through kernel mapping $K(x,y)=\Phi(x)\Phi(y)$. Then, KPCA is applied. The proposed method is more effective on feature extraction than linear method.

Given a set of sample X , feature space of samples is obtained by kernel mapping Φ which is sorted as sample matrix in column-vector $\Psi=[q_1,q_2,\dots,q_N]$, $q_i=\Phi(x_i)$, $i=1,\dots,N$. Suppose that samples' mean are zero, whose covariance matrix is

$$C = \frac{1}{N} \sum_{i=1}^N q_i q_i^T = \frac{1}{N} \Psi \Psi^T .$$

Since we cannot obtain the explicit knowledge about formation of samples, hence, we assume that feature vector is expressed by sample's linear combination:

$v = \sum_{i=1}^N a_i q_i = \Psi a$. However, if sample matrix Ψ is singular, we can not obtain feature

vector in the null space of sample. Therefore, the feature decomposition of covariance matrix can be stated as: $C v = p v = \frac{1}{N} \Psi \Psi^T \Psi a = p \Psi a$, where p is feature value. Multiply above formula by sample matrix Ψ^T :

$$\frac{1}{N} \Psi^T \Psi \Psi^T \Psi a = \frac{1}{N} K K a = p K a = p \Psi^T \Psi a \tag{1}$$

Where $K = [k(x_i, x_j)]_{i,j=1}^N$, herein, (1) can be simplified as $K a = p a$.

4 Test and Simulation

Based on the theory mentioned above, combining the Curvelet with KPCA, the procedure of algorithm is briefed as:

- 1) Image pre-processing, where images are normalized to the size of 256×256 .
- 2) Curvelet transform on images, where Curvelet transform is applied on each image in the face database to extract the first layer of Curvelet coefficients $C\{1\}$ which are

arranged into a column vector and composed to be the Curvelet characteristics T .(scale=3,angle=8).

3) KCPA, where polynomial kernel principal component $K(x_i, x_j)=(x_i \cdot x_j+1)^d$ is used to extract the main component of the Curvelet feature T as the recognition feature of the face (train_features).

4) Recognition, where the same approach is employed that the feature of tested images are extracted. Furthermore the nearest method is adopted for classification.

The face recognition experiments are carried out in Matlab7.6.0, with a dual-CORE 2.0 GHZ processor and 2GB RAM. Herein, the JAFFE database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models, while the FERET database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images. We randomly select 20 people and 10 photos in each.

We also do experiments under the ORL face database that consists of 400 images where each of 40 distinct subjects comes along with ten different images, including expression variation, slight posture variation. Different values of kernel function parameter d are trialed where five images are selected as training samples and another five images as test samples. The kernel parameter changes over the range of $0.01 \sim 1$.

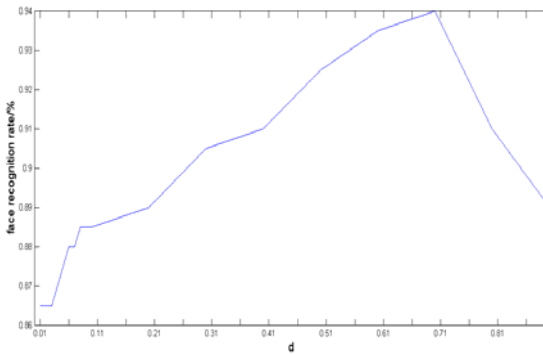


Fig. 2. Recognition rate with different parameters in the ORL database

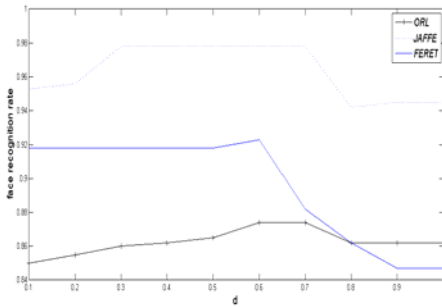
It is demonstrated that the better recognition rate is achieved when the kernel function parameter d is set between 0.1 and 1 as shown in Fig.2. The detail of relationship between d and recognition in ORL, JAFFE and FERET face database is shown in Fig.3. We choose number of samples as 3, 4, 5, and 6, respectively. Thus the best result is generated at around $d=0.7$ which is our optimal test parameter.

In order to further validate the effectiveness of the proposed algorithm, a series of comparative experiments are carried out in the ORL, JAFFE, FERET face database in terms of proposed algorithm and the algorithms mentioned in the literature [8][11] (The training samples are chosen as 5 images per person). The comparison result is showing in the table 1.

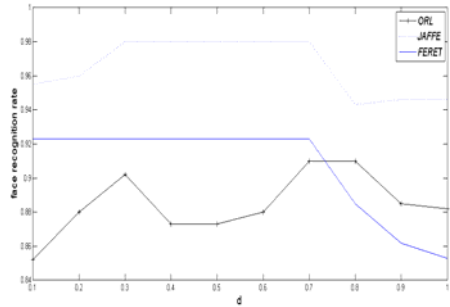
Table 1. Comparison of test results

	Feature dimension	Recognition rate (%) (ORL)	Recognition rate (%) (JAFFE)	Recognition rate (%) (FERET)
Curvelet coarse	7225	86.50	89.45	84.44
Curvelet coarse +PCA	6500	90.77	92.51	89.95
Wavelet +KPCA	199	92.45	94.68	90.72
KPCA	106	87.35	81.08	83.25
This Paper	164	94.05	97.60	94.82

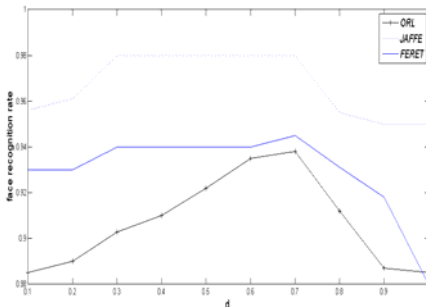
In table 1 we can find that the recognition feature dimension extracted by the coarse layer of the Curvelet and the Curvelet +PCA is significantly higher than the proposed algorithm. Compared with wavelet algorithm, in all three databases, the proposed algorithm is overall better than the algorithm combined the wavelet with KPCA under all the indices.



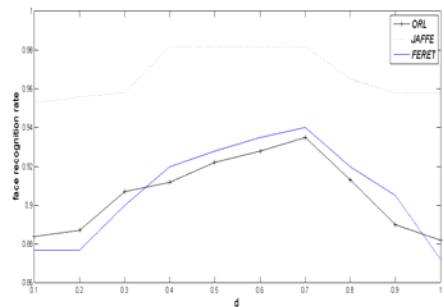
(a) Three training samples



(b) Four training samples



(c) Five training samples



(d) Six training samples

Fig. 3. Recognition rate with different parameters and training samples in three different face databases

5 Conclusions

This paper proposes a Curvelet transform based face recognition method through obtaining Curvelet coefficients with kernel principal component analysis. Furthermore, we do recognition experiments with varied parameters of d in three face database. Compared with wavelet transform methods, the proposed algorithm not only increases the recognition rate but also significantly reduces the dimension of extracted feature. In order to achieve better performances, such as recognition rate, we will further focus on the kernel function and the classifier design in the later work.

References

1. Jiao, L.-C., Gong, M.-G., Wang, S., et al.: *Advances in Natural Computation, Machine Learning and Image Understanding*, pp. 224–243. Xidian University Press, Xi'an (2008)
2. Nikam, S.B., Agarwal, S.: Curvelet-based fingerprint anti-spoofing. *Signal, Image and Video Processing* 4(1), 75–87 (2010)
3. Lakshmi, A., Rakshit, S.: New Curvelet Features for Image Indexing and Retrieval. *Computer Networks and Intelligent Computing* 157(9), 492–501 (2011)
4. Wu, X.-X., Zhao, J.-Y., Shen, M.-X.: Face Processing in Human-computer Interaction Using Curvelet Analysis. *Image and Graphics* 15(9), 1309–1317 (2010)
5. Amira, A., Farrell, P.: An Automatic Face Recognition System Based on Wavelet Transforms. In: *IEEE International Symposium on Circuits and Systems*, Kobe, Japan, May 23–26, pp. 6252–6255 (2005)
6. Liu, C.-J., Wechsler, H.: Independent Component Analysis of Gabor Features for Face Recognition. *IEEE Transactions on Neural Networks (S1045-9227)* 14(4), 919–928 (2003)
7. Su, H.-T., Zhao, R.-C.: Face Recognition Based on Wavelet Transform and Multiple Classifier. *Computer Applications* 22(8), 25–27 (2002)
8. Rahman, S., Naim, S.M., Al Farooq, A., Islam, M.M.: Curvelet Texture Based Face Recognition Using Principal Component Analysis. In: *Proceedings of 13th International Conference on Computer and Information Technology (ICCIT 2010)*, Dhaka, Bangladesh, December 23–25, pp. 45–50 (2010)
9. Mandal, T., Jonathan Wu, Q.M., Yuan, Y.: Curvelet Based Face Recognition via Dimension Reduction. *Signal Processing (S0165-1684)* 89(12), 2345–2353 (2009)
10. Xu, X.-B., Zhang, D.-Y., Zhang, X.-M., et al.: An Algorithm for Multimodal Biometric Recognition Based on Feature Level and the Second-Generation Curvelet Transform. *Journal of Xi'an Jiaotong University* 43(10), 32–36 (2009)
11. Li, W.-H., Gong, W.-G., Cheng, W.-M., et al.: Method Based on Wavelet Multiresolution Analysis and KPCA for Face Recognition. *Computer Applications* 25(10), 2339–2341 (2005)
12. Li, W., Gong, W., Liang, Y., Chen, W.-M.: Feature Selection Based on KPCA, SVM and GSFS for Face Recognition. In: Singh, S., Singh, M., Apte, C., Perner, P. (eds.) *ICAPR 2005*. LNCS, vol. 3687, pp. 344–350. Springer, Heidelberg (2005)
13. Wu, X.-X., Zhao, J.-Y.: Curvelet Feature Extraction for Face Recognition and Facial Expression Recognition. In: *2010 Sixth International Conference on Natural Computation (ICNC 2010)*, Yantai, Shandong, China, August 10–12, pp. 1212–1216 (2010)

Marine Spill Oil SAR Image Segmentation Based on Maximum Entropy and CV Model

Yang Ji^{1,2}, Yiquan Wu¹, and Yi Shen¹

¹ College of Electronic and Information Engineering,
Nanjing University of Aeronautics and Astronautics, 210016 Nanjing, China

² Open Fund of the Key Laboratory of Ocean Circulation and Waves,
Chinese Academy of Sciences, 266071 Qingdao, China
xiaobainj@sina.com

Abstract. To solve the problem that the accuracy of SAR image segmentation is not high enough in the marine spill oil detection, a segmentation method of marine spill oil images based on maximum entropy and CV model is proposed in this paper. Firstly, the multilevel thresholding algorithm based on maximum entropy is used to make a coarse segmentation for marine spill oil images. The obtained spill oil region and coarse contour provide local region and initial contour for CV model, respectively, to reduce the scene complexity of CV model and its sensitivity to initial situation. That is CV model is utilized to subdivide the local area. Lots of experimental results show that the proposed segmentation method of marine spill oil SAR images not only enables the dispense with initial condition but also ensures accurate segmentation contour and efficient operation.

Keywords: marine spill oil detection, synthetic aperture radar remote sensing image, image segmentation, maximum entropy, CV model.

1 Introduction

Marine oil spill is one of the most common ocean pollutants [1], which can be caused by the wretched weather, illegal drainage of the contaminated oil and the accidents happened on the operation platform. In order to relieve the damage to the eco-environment caused by the onshore oil spilling accident, maritime countries from all over the world usually adopt all kinds of technology methods such as offshore monitoring stations, ships, auto observing platforms, aviation and satellite remote sensing to carry out the stereoscopic supervision on marine oil spill [2]. Synthetic Aperture Radar SAR remote sensing, a full time detecting technology can be used to continuously detect the changes of large area marine oil spill [3], especially in cracking down the malicious illegal excretion of contaminated oil [4]. The detecting technology, key to SAR remote sensing images of marine oil spill, consists of many parts such as preconditioning, segmentation, feature extraction and classification [5]. People have already made some researches on it. LIANG Xiao-yi[6], etc, research the selection of texture characteristics in classifying oil spill SAR images. Some discussion about oil spill identification based on texture feature and artificial neural network are

made in references [7-8]. Ma Teng-bo and Wang Si-yuan[9] proposed an oil detection algorithm based on edge detection. In the oil spill monitoring technology, image segmentation method is one of the key steps and directly impacts the final results [10]. Zou Ya-rong[11], etc, study on segmentation of SAR image for oil spilled. But right now the special research of oil spill SAR image segmentation is less and the accuracy is not high enough. Oil spills appear as dark areas in the SAR images because oil dampens the Bragg waves of the area surface and reduces the radar backscattering coefficient. Therefore, marine spill oil SAR image segmentation is essentially extracted the dark area in the images [12]. Hence, method based on edge, method based on region and method based on the combination of edge and region can be used for marine spill oil detection. The last method can use the advantages of the other two methods, but usually need choose random selection of initial conditions and the efficiency is not high.

Thresholding is one of the most important techniques for performing image segmentation which is based on the region. It is generally simple and computationally efficient. The main objective is to determine the optimal threshold. Maximum Shannon entropy is one of the most concern threshold selection methods and better segmentation effect is expected. But the thresholding method only using gray information and there are many speckle noises in the SAR image, gray information can not reflect the structure information of the images very well. Therefore, the accuracy certainly will not high and it can hardly use in the SAR images alone. On the other hand, as the representative of the edge segmentation, under the inspiration of Mumford-Shah function [13], Chan and Vese[14] proposed the active contours without edges, namely Chan-Vese model. CV model can give satisfactory results of weak edge image; it is a classical and practical model. Because of the only uses the regional mean value to define the boundary in the mode, it is difficult to use in complex scene, and still exists the defects of low efficiency and sensitivity of the initial conditions. If we can import the thresholding based on maximum entropy and CV model into the field of marine spill oil detection and consider combining them further, the segmentation precision and operation efficiency are hopefully increased in marine spill oil detection.

Some researches of the application of multilevel thresholding based on maximum entropy and CV model in the marine spill oil segmentation are made in this paper. On this basis, a marine spill oil detection method based on maximum entropy and CV model is proposed. Firstly, the multilevel thresholding based on maximum entropy is used to make a coarse segmentation. The obtained spill oil region and coarse contour provide local region and initial contour for CV model, respectively. Then, CV model is utilized to subdivide the local area.

2 Marine Spill Oil SAR Image Coarse Segmentation Based on Maximum Entropy

When the quality of image is good and the requirement of segmentation accuracy is not high, the only use of thresholding can achieve satisfactory results. Oil spills appear as dark areas in SAR images because oil dampens the Bragg waves of the sea surface and reduces the radar backscattering coefficient. The critical part of the oil spill detection is to distinguish oil spills from other natural phenomena. They can be

due to low wind speed areas, internal waves, biogenic films, algal blooms, grease ice, wind front areas, areas sheltered by land, rain cells, current shear zones and upwelling zones. If the oil spill areas have apparent contrast with other areas, the single level thresholding can be employed. But actually, there are many multiplicative speckle noises in the SAR images which will greatly affect segmentation results. Even take speckle reducing as the pre-processing; the desired results can't appear simply through the thresholding. To improve the accuracy, a segmentation method of marine spill oil images based on maximum entropy and CV model is proposed in this paper. Firstly, the multilevel thresholding algorithm based on maximum entropy is used to make a coarse segmentation for marine spill oil images. The obtained spill oil region and coarse contour provide local region and initial contour for CV model, and then CV model is utilized to subdivide the local area.

2.1 Definition of Maximum Entropy and the Selection of Signal Threshold

We assume the scale of image f is $M \times N$, its gray level is $0, 1, \dots, L-1$, let p_i express the probability of level i which satisfy $\sum_{i=1}^{L-1} p_i = 1$. From this distribution we use threshold t

to derive two probability distributions, one for the object (class O) and another for the background (class B), that is $C_O = \{(m, n) | f(m, n) = 0, 1, \dots, t\}$ and

$$C_B = \{(m, n) | f(m, n) = t+1, t+2, \dots, L-1\} . \text{ Let } P_t = \sum_{i=0}^t p_i , \quad h_t = -\sum_{i=0}^t p_i \ln p_i ,$$

$0 \leq t \leq L-1$, $1-P_t = \sum_{i=t+1}^{L-1} p_i$, the probability distributions of the object and back-

ground classes are given by $\left\{ \frac{p_0}{P_t}, \frac{p_1}{P_t}, \dots, \frac{p_t}{P_t} \right\}$ and $\left\{ \frac{p_{t+1}}{1-P_t}, \frac{p_{t+2}}{1-P_t}, \dots, \frac{p_{L-1}}{1-P_t} \right\}$, obvious-

ly, $\sum_{i=0}^t \frac{p_i}{P_t} = 1$, $\sum_{i=t+1}^{L-1} \frac{p_i}{1-P_t} = 1$. In the situation of single threshold value, the entropy for each distribution is defined as

$$H_O(t) = -\sum_{i=0}^t \frac{p_i}{P_t} \ln \frac{p_i}{P_t} = \ln P_t + \frac{h_t}{P_t} . \tag{1}$$

$$H_B(t) = -\sum_{i=t+1}^{L-1} \frac{p_i}{1-P_t} \ln \frac{p_i}{1-P_t} = \ln(1-P_t) + \frac{H-h_t}{1-P_t} . \tag{2}$$

The total entropy $H(t)$ is formulated as the sum of each entropy

$$H(t) = H_O(t) + H_B(t) = \ln P_t(1-P_t) + \frac{h_t}{P_t} + \frac{H-h_t}{1-P_t} . \tag{3}$$

When $H(t)$ is maximized, the luminance level t is the optimum threshold value

$$t^* = \underset{0 \leq t < L-1}{\text{Arg max}} \{H(t)\}. \tag{4}$$

2.2 Selection of Multiple Thresholds Based on Maximum Entropy

The selection of one threshold based on the maximum entropy above can be extended to the multiple situations. Now use $n-1$ thresholds t_1, t_2, \dots, t_{n-1} ($0=t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n=L-1$) to divide the pixel of image f into n class $C_k(k=1, 2, \dots, n)$. Let $T=(t_1, t_2, \dots, t_{n-1})$, then the maximum entropy $H(T)$ of the n class is

$$\begin{aligned} H(T) &= \ln \left(\sum_{i=0}^{t_1} p_i \right) + \ln \left(\sum_{i=t_1+1}^{t_2} p_i \right) + \dots + \ln \left(\sum_{i=t_{n-1}+1}^{L-1} p_i \right) - \\ &\quad \frac{\sum_{i=0}^{t_1} p_i \ln p_i}{\sum_{i=0}^{t_1} p_i} - \frac{\sum_{i=t_1+1}^{t_2} p_i \ln p_i}{\sum_{i=t_1+1}^{t_2} p_i} - \dots - \frac{\sum_{i=t_{n-1}+1}^{L-1} p_i \ln p_i}{\sum_{i=t_{n-1}+1}^{L-1} p_i} \\ &= -\frac{w(0, t_1)}{u(0, t_1)} + \ln u(0, t_1) - \frac{w(t_1+1, t_2)}{u(t_1+1, t_2)} + \ln u(t_1+1, t_2) \\ &\quad - \dots - \frac{w(t_{n-1}+1, t_n)}{u(t_{n-1}+1, t_n)} + \ln u(t_{n-1}+1, t_n). \end{aligned} \tag{5}$$

where $w(0, t_1) = \sum_{i=0}^{t_1} p_i \ln p_i$, $u(0, t_1) = \sum_{i=0}^{t_1} p_i$,

$$w(t_{k-1}+1, t_k) = \sum_{i=t_{k-1}+1}^{t_k} p_i \ln p_i, u(t_{k-1}+1, t_k) = \sum_{i=t_{k-1}+1}^{t_k} p_i, k = 2, 3, \dots, n.$$

In order to reduce the time of looking for $n-1$ optimal threshold values, the recursive formula for computing $w(t_{k-1}+1, t_k)$ and $u(t_{k-1}+1, t_k)$ are given as follows

$$w(0, 0) = 0, \quad u(0, 0) = 0$$

$$\begin{cases} w(0, t) = \sum_{i=0}^t p_i \ln p_i = w(0, t-1) + p_t \ln p_t \\ u(0, t) = \sum_{i=0}^t p_i = u(0, t-1) + p_t \end{cases} \quad t = 1, 2, \dots, L-1. \tag{6}$$

$$\begin{cases} w(t_{k-1}+1, t_k) = w(0, t_k) - w(0, t_{k-1}) \\ u(t_{k-1}+1, t_k) = u(0, t_k) - u(0, t_{k-1}) \end{cases} \quad k = 2, 3, \dots, n. \tag{7}$$

When $H(T)$ is maximized, the $n-1$ luminance levels $T=(t_1, t_2, \dots, t_{n-1})$ is the optimum thresholds

$$(t_1^*, t_2^*, \dots, t_{n-1}^*) = \underset{0 \leq t_1 < t_2 < \dots < t_{n-1} \leq L-1}{\text{Arg max}} \{H(T)\}. \tag{8}$$

The employ of recursive formulae (6) and (7) can reduce the computational complexity and improve the operation efficiency. In order to further shorten the computation time, the optimal thresholds are searched by the LZM PSO algorithm [15] in this paper.

3 Subdivision of Marine Spill Oil Images Based on CV Model

Let $I(C)$ and $O(C)$ represent the inside and outside of evolving curve C . Assume that the image z is formed by two regions of approximatively piecewise-constant intensities. CV model can be achieved by extending the formula as follows

$$F_1(C) + F_2(C) = \int_{I(C)} |z(x) - c_1|^2 dx + \int_{O(C)} |z(x) - c_2|^2 dx \tag{9}$$

where z is the gray value of the image, and the constants c_1, c_2 depending on C , are the averages inside C and respectively outside C . Therefore, the energy functional $F(c_1, c_2, C)$, defined by

$$F(c_1, c_2, C) = \mu L_e(C) + \nu A(I(C)) + \lambda_1 \int_{I(C)} |z(x) - c_1|^2 + \lambda_2 \int_{O(C)} |z(x) - c_2|^2. \tag{10}$$

where $\mu \geq 0, \nu \geq 0, \lambda_1, \lambda_2 \geq 0$ are fixed parameters. In almost all our numerical calculations, we fix $\lambda_1 = \lambda_2 = 1$ and $\nu = 0$. $L_e(C)$ and $A(I(C))$ represent the length of the curve C and the area of the region inside C , respectively. The last two members play the leading role, namely the sum of bias squares about inside and outside with respect to their own average gray.

Map [10] into the level set formulation, we get

$$F(c_1, c_2, \phi) = \mu \int_{\Omega} \delta(\phi) |\nabla \phi| dx + \nu \int_{\Omega} H(\phi) dx + \lambda_1 \int_{\Omega} |z(x) - c_1|^2 H(\phi) dx + \lambda_2 \int_{\Omega} |z(x) - c_2|^2 (1 - H(\phi)) dx. \tag{11}$$

where δ is the Dirac function, H is the Heaviside function, ϕ is the level set function. Usually take the H function generated by the contour C . Keeping ϕ fixed and minimizing the energy function with respect to the constants c_1 and c_2 , it is easy to express these constants function of ϕ by

$$c_1 = \frac{\int z(x) H(\phi) dx}{\int H(\phi) dx}, \quad c_2 = \frac{\int z(x) (1 - H(\phi)) dx}{\int (1 - H(\phi)) dx}. \tag{12}$$

Keeping c_1 and c_2 fixed, and minimizing the energy function with respect to ϕ , we deduce the associated Euler-Lagrange equation for ϕ

$$\delta(\phi) \left[\mu \nabla \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - \nu - \lambda_1 (z - c_1)^2 + \lambda_2 (z - c_2)^2 \right] = 0. \tag{13}$$

4 Experiment Results and Analysis

Experiment three methods (multilevel thresholding based on maximum entropy, CV model, the method of combination of multilevel thresholding based on maximum entropy and CV model, notes for method 1, method 2, method 3, respectively) on lots of real marine spill oil images. Parameters of the methods are generally chosen as follows: time step is 0.5, the space step is 1, $\mu=0.01 \times 255^2$, $\lambda_1=\lambda_2=1, \nu=0$. In the method of multilevel thresholding, we choose two threshold values which divide the image into three categories. In the segmentation results, the minimal gray level is the oil spill area, the rest can be considered as the background.

From Fig.1 to Fig.3, (a) represents the original image, and (b)~(d) represent the results of method 1~method 3, respectively. There exists lots of image speckle noise in Fig.1, and the contrast of image is relatively low. Maximum entropy criterion is used only to get a high false alarm, almost ineffective, while method based on CV model can get a outline, but also has some false alarm. The combination of multi-level threshold segmentation by maximum entropy criterion and CV model method can get an oil spilled region with a more accurate result and a better anti-noise ability. The speckle image of spilled oil area is stripe shaped in Fig.2. What's more, the multi-level threshold segmentation by maximum entropy criterion and CV model method are both weak in anti-nose themselves, thus can't get an accurate oil outline. However, the combination of these two methods can get a precise outline and have ideal anti-noise ability. The outline showed in Fig. 3 is consisted by an infinite variety of forms which are either smooth or continuous, concave-convex transited, or anvil and sharp pointed. Both these two methods can be used to segregate the image, but there exists some false alarm and the edges got are rough. The combination of multilevel thresholding by maximum entropy criterion and CV model method can be used to get an oil spilled region with a more accurate segmentation and a stronger anti-noise ability.

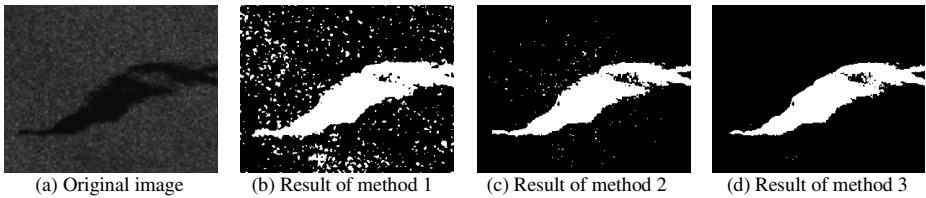


Fig. 1. Comparison of experimental results (the first group)

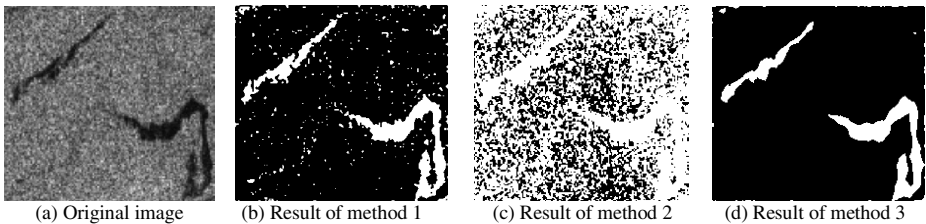


Fig. 2. Comparison of experimental results (the second group)

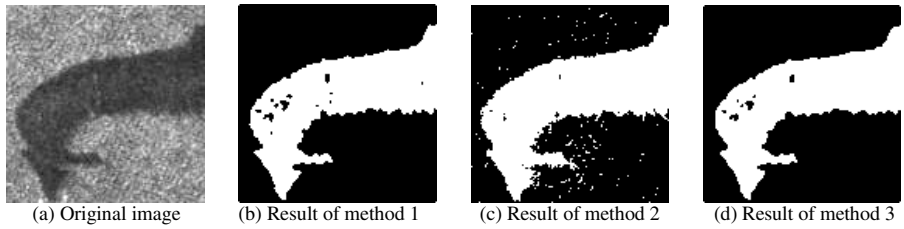


Fig. 3. Comparison of experimental results (the third group)

Then quantitatively evaluate are made according to normalized log likelihood and variance of the ratio image, noted for D and RI_{var} [16]. The smaller the two indexes mean the better the segmental performance. The segmental performance of the above three methods are listed in Table 1.

Table 1. Performance comparison of three methods

RI_{var} / D	<i>Method 1</i>	<i>Method 2</i>	<i>Method 3</i>
Group Index			
1	0.3693/0.5182	0.3400/0.4532	0.3315/0.4137
2	0.2771/0.2708	0.3156/0.7294	0.2286/0.1572
3	0.3446/0.4517	0.3651/0.4680	0.3426/0.4430

The experimental environment is Intel Celeration(R) CPU 2.66GHz/512MB memory/Matlab 7.0. The operation times and iterations of the three methods are listed in Table 2 noted as t /number and are counted in second and unit, respectively. From Table 2, we know that method 1 doesn't need iterations and run fastest; method 2 need many iterations and long operation time; method 3 greatly improve the convergence rate of method 2(CV model), the iterations are only one percent of the originals and the running times are significantly shortened.

Table 2. The operation times and iterations of three methods

$t / number$	<i>Method 1</i>	<i>Method 2</i>	<i>Method 3</i>
Group Index			
1	0.210/non	269.342/1435	4.971/15
2	0.198/non	546.219/2000	3.592/10
3	0.179/non	55.254/943	0.875/3

5 Conclusion

To solve the problem that the accuracy of SAR image segmentation is not high enough in the marine spill oil detection, after making full consideration of

thresholding based on the maximum entropy and CV model, a segmentation method of marine spill oil images based on maximum entropy and CV model is proposed in this paper. The method combines the global property of maximum entropy with the local optimal of the CV model, which can solve the problems of low segmentation accuracy and operation efficiency. Experiment results of real marine spill oil images show that the method based on maximum entropy and CV model not only has advantages over the segmental results, but also has less operation times. Therefore, it is a marine spill oil images thresholding method of high practical value. Then the further extraction of the spill regional characteristics can be used for automatic identification of oil spill which are our next research work.

Acknowledgments. Foundation: Open Fund of the Key Laboratory of Ocean Circulation and Waves, Chinese Academy of Sciences (No.KLOCAW1110); Open Foundation of the Key Laboratory of Marine Spill Oil Identification and Damage Assessment Technology (No.201112); Open Foundation of the Key Laboratory of Underwater Acoustic Communication and Marine Information Technology (Xiamen University), Ministry of Education, Xiamen (Grant No.201101); National Natural Science Foundation of China (No.60872065).

References

1. Wang, G.W., Zhang, Y.Z., Lin, H.: A Study of Oil Spill Detection Using ASAR Images. *Acta Oceanologica Sinica* 28, 32–37 (2009)
2. Lu, J.: Marine Oil Spill Detection, Statistics and Mapping with ERS SAR Imagery in South-east Asia. *International Journal Remote Sensing* 24, 3013–3032 (2003)
3. Fanny, G.A., Grgoire, M., Fabrice, C., Ren, G.: Operational Oil-slick Characterization by SAR Imagery and Synergistic Data. *IEEE Journal of Oceanic Engineering* 30, 487–495 (2005)
4. Camilla, B., Solberg, A.H.S.: Oil Spill Detection by Satellite Remote Sensing. *Remote Sensing of Environment* 95, 1–13 (2005)
5. Migliaccio, M., Gambardella, A., Tranfaglia, M.: SAR Polarimetry to Observe Oil Spills. *IEEE Transactions Geoscience Remote Sensing* 45, 506–511 (2007)
6. Liang, X.Y., Zhang, J., Meng, J.: Selection of Texture Characteristics in Classifying Oil Spill SAR Images. *Advance in Marine Science* 25, 346–354 (2007)
7. Topouzelis, K., Karathanassi, V., Pavlakis, P., Rokos, D.: Detection and Discrimination between Oil Spills and Look-alike Phenomena through Neural Networks. *ISPRS Journal of Photogrammetry & Remote Sensing* 62, 264–270 (2007)
8. Shi, L.J., Zhao, C.F., Liu, P.: Oil Spill Identification Based on Texture Feature and Artificial neural Network. *Periodical of Ocean University of China* 39, 1269–1274 (2009)
9. Ma, T.B., Wang, S.Y.: Oil Detection Algorithm Based on Edge Detection. *Journal of Remote Sensing* 13, 1087–1091 (2009)
10. Lena, C., Chen, C.M., Tang, J.C.: An Automatic Detection of Oil Spills in SAR Images by Using Image Segmentation Approach. In: *IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, pp. 1021–1024 (2005)
11. Zou, Y.R., Wang, H., Zhu, H.T., Chen, G.M., Song, X.A.: Study on Segmentation of SAR Image for Oil Spilled at Sea. *Marine Environment Science* 28, 313–315 (2009)

12. Solberg, A.H.S., Brekke, C.: Oil Spill Detection in Radarsat and Envisat SAR Images. *IEEE Transactions Geoscience Remote Sensing* 45, 746–755 (2007)
13. Mumford, D., Shan, J.: Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems. *Communications and Pure and Applied Mathematics* 42, 577–685 (1989)
14. Chan, T., Vese, L.: Active Contours without Edges. *IEEE Transactions on Image Processing* 10, 266–277 (2001)
15. Liu, D.H., Yuan, S.C., Lan, Y.: Method of Particle Swarm Optimization Based on the Chaos Map. *Journal of Xidian University* 37, 764–769 (2010)
16. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. *International Journal of Computer Vision* 22, 61–79 (1997)

Multilevel Thresholding Based on Exponent Gray Entropy and Niche Chaotic Particle Swarm Optimization

Yi Shen^{1,2}, Yiquan Wu², and Yang Ji²

¹ Key Laboratory of Nondestructive Testing (Nanchang Hangkong University),
Ministry of Education, 330063 Nanchang, China

² College of Electronic and Information Engineering,
Nanjing University of Aeronautics and Astronautics, 210016 Nanjing, China
yanghe_shenyi@163.com

Abstract. The method of threshold selection based on maximal Shannon entropy or exponent entropy only depend on the probability information from gray image histogram, and don't immediately consider the uniformity of the gray scale within the cluster. Considering these facts, thresholding based on exponent gray entropy is proposed. Firstly, exponent gray entropy is defined and the method of single threshold selection is given. Then, the method is extended to multilevel thresholding. Furthermore, the niche chaotic mutation particle swarm optimization algorithm is adopted to find the best multi-threshold. Many experimental results show that, compared with multilevel thresholding based on maximal entropy and particle swarm optimization, the proposed segmentation method has less operation times and segmented images of the suggested method are more accurate in edge and texture.

Keywords: threshold selection, exponent gray entropy, multi-threshold, chaotic mutation, niche particle swarm optimization.

1 Introduction

As the basic concern in the computer vision and image processing, image segmentation is the foundation of the target detection, character extraction and identification. Thresholding is a commonly-used, simple and effective image segmentation method. The heart of the problem is to select an appropriate threshold quickly and accurately for the segmentation[1]. The scholars home and abroad have done a lot of research on this issue[2,3], and have put forward many kinds of thresholding methods. Among them, the maximum Shannon entropy thresholding is one of the most important threshold selection methods[4]. But the problem of computationally prohibitive cannot be ignored when the maximum Shannon entropy criterion is applied to multilevel thresholding[5,6]. To deal with this problem, people proposed simulated annealing algorithm[7], iterated conditional modes algorithm[8], genetic algorithm[9-12] and particle swarm optimization (PSO)[13-15] for multilevel thresholding. Although Shannon entropy defined by logarithmic is effectively used to measure information uncertainly, there exists problem of undefined value and zero value. In 1988, Pal S K

and Pal N R proposed a new definition of classical entropy based on exponential behavior which overcome the lack of Shannon entropy[16]. But this method only depends on the probability information from gray image histogram, and does not directly consider the uniformity of the gray scale within the cluster. In addition, PSO is simple, easy to implement and the parameters to be determined are less, while it is easy trapping into partial extreme and the accuracy of the searching is lower. For this purpose, this paper will use niche chaotic mutation particle swarm optimization (NCPSO)[17] to search the optimum thresholds.

Based on the above analysis, this paper puts forward a method of multilevel thresholding which is based on exponent gray entropy and NCPSO. Exponent gray entropy reflects the uniformity of the gray scale immediately within the cluster. And NCPSO can find the optimum thresholds fast and accurate. At last in the experiment results and analysis, the multilevel thresholding segmentation outcome and running time of the method in this paper will be given, and also the method will be compared with multi-threshold selection based on maximal Shannon entropy and PSO[13].

2 Thresholding Based on Exponent Gray Entropy

We assume the gray scale $f(m, n)$ of a $M \times N$ image is picked in $0, 1, \dots, L-1$, and let $h(i)$ be the number of pixels with gray value i . Then the occurrence frequency of gray level i is $p(i) = h(i)/(M \times N)$, where $p(i)$ satisfies $\sum_{i=1}^{L-1} p_i = 1$. Now we divide the image into object class $C_o = \{(m, n) | f(m, n) = 0, 1, \dots, t\}$ and background class $C_b = \{(m, n) | f(m, n) = t + 1, t + 2, \dots, L - 1\}$ by threshold t . Let

$$P_{m,n} = \begin{cases} \frac{f(m,n)}{\sum_{(x,y) \in C_o} f(x,y)}, & (m,n) \in C_o \\ \frac{f(m,n)}{\sum_{(x,y) \in C_b} f(x,y)}, & (m,n) \in C_b \end{cases} \quad (1)$$

obviously, $\sum_{(m,n) \in C_o} P_{m,n} = 1$. Then the exponent gray entropy of the object class is defined as:

$$\begin{aligned} H_o(t) &= \sum_{(m,n) \in C_o} P_{m,n} e^{1-P_{m,n}} = \sum_{(m,n) \in C_o} \frac{f(m,n)}{\sum_{(x,y) \in C_o} f(x,y)} e^{1 - \frac{f(m,n)}{\sum_{(x,y) \in C_o} f(x,y)}} \\ &= \sum_{i=0}^t h(i) \frac{i}{\sum_{i'=0}^t h(i') i'} e^{1 - \frac{i}{\sum_{i'=0}^t h(i') i'}} \end{aligned} \quad (2)$$

and the exponent gray entropy of the background class is defined as:

$$\begin{aligned}
 H_b(t) &= \sum_{(m,n) \in C_b} p_{m,n} e^{1-p_{m,n}} = \sum_{(m,n) \in C_b} \frac{f(m,n)}{\sum_{(x,y) \in C_b} f(x,y)} e^{1-\frac{f(m,n)}{\sum_{(x,y) \in C_b} f(x,y)}} \\
 &= \sum_{i=i'+1}^{L-1} h(i) \frac{i}{\sum_{i'=i'+1}^{L-1} h(i')i'} e^{1-\frac{i}{\sum_{i'=i'+1}^{L-1} h(i')i'}}.
 \end{aligned}
 \tag{3}$$

So the total exponent gray entropy is:

$$H(t) = H_o(t) + H_b(t) = \sum_{i=0}^t h(i) \frac{i}{u_o(t)} e^{1-\frac{i}{u_o(t)}} + \sum_{i=i'+1}^{L-1} h(i) \frac{i}{u_b(t)} e^{1-\frac{i}{u_b(t)}}.
 \tag{4}$$

where $u_o(t) = \sum_{i=0}^t h(i)i$, $u_b(t) = \sum_{i=i'+1}^{L-1} h(i)i$. The exponent gray entropy is larger; the difference of the pixels' gray scale within the cluster is smaller. The segmentation result is best when $H(t)$ reaches maximum. Therefore the optimal threshold of the maximum exponent gray entropy is:

$$t^* = \underset{0 \leq t \leq L-1}{\text{Arg max}} \{H(t)\}.
 \tag{5}$$

3 Multilevel Thresholding Based on Exponent Gray Entropy and NCPSO

3.1 Multilevel Thresholding Based on Exponent Gray Entropy

We can extend (4) to get multi-threshold selection:

$$\begin{aligned}
 H(t_1, t_2, \dots, t_n) &= \sum_{i=0}^{t_1} h(i) \frac{i}{u(0, t_1)} e^{1-\frac{i}{u(0, t_1)}} + \sum_{i=t_1+1}^{t_2} h(i) \frac{i}{u(t_1+1, t_2)} e^{1-\frac{i}{u(t_1+1, t_2)}} + \dots \\
 &+ \sum_{i=t_{n-1}+1}^{t_n} h(i) \frac{i}{u(t_{n-1}+1, t_n)} e^{1-\frac{i}{u(t_{n-1}+1, t_n)}} + \sum_{i=t_n+1}^{L-1} h(i) \frac{i}{u(t_n+1, L-1)} e^{1-\frac{i}{u(t_n+1, L-1)}}.
 \end{aligned}
 \tag{6}$$

where t_1, t_2, \dots, t_n are thresholds, $0 \leq t_1 < t_2 < \dots < t_{n-1} < t_n < L-1$, $u(0, t_1) = \sum_{i=0}^{t_1} h(i)i$;

$u(t_{k-1}+1, t_k) = \sum_{i=t_{k-1}+1}^{t_k} h(i)i$, $k = 2, 3, \dots, n$; $u(t_n+1, L-1) = \sum_{i=t_n+1}^{L-1} h(i)i$. Therefore the optimal multi-threshold is:

$$(T_1^*, T_2^*, \dots, T_n^*) = \underset{0 \leq t_1 < t_2 < \dots < t_{n-1} \leq L-1}{\text{Arg max}} \{H(t_1, t_2, \dots, t_n)\}.
 \tag{7}$$

3.2 Algorithm Based on NCPSO

In order to reduce the computation, we use NCPSO to search the optimum multi-threshold.

Assume in the N -dimensional space, every particle l has the position $X_l = (X_{l1}, X_{l2}, \dots, X_{ln})$ and the speed $V_l = (V_{l1}, V_{l2}, \dots, V_{ln})$. X_l represents the answer of the question and the corresponding objective function value is used for evaluating the advantages and disadvantages of particle fitness; V_l represents the speed of the particle moving from one place to another. The first step is the initialization of particle swarm, and then it finds the optimal answer by iterative equation. Assuming in the iteration of m , the optimal answer of the particle l is $\mathbf{pbest}_l(m)$, called individual extreme; the optimal answer of the particle swarm is $\mathbf{gbest}(m)$, called global extreme. At the $m+1$ moment it updates its speed by the following formula:

$$V_l(m+1) = wV_l(m) + c_1r_1[\mathbf{pbest}_l(m) - X_l(m)] + c_2r_2[\mathbf{gbest}(m) - X_l(m)]. \tag{8}$$

Then it moves to the next place by the speed of $V_l(m+1)$, as:

$$X_l(m+1) = X_l(m) + V_l(m+1). \tag{9}$$

where m stands for iterative time, $c_1 = c_2 = 2$ is the learning factor; r_1, r_2 are the random number between 0 and 1; inertia factor is $w = w_{\max} - (w_{\max} - w_{\min})m / m_{\max}$, where m_{\max} stands for whole iterative time, w_{\max} and w_{\min} separately stands for the maximum and minimum inertia factor (in this paper $m_{\max} = 100, w_{\max} = 0.95, w_{\min} = 0.4$). In the iteration and updating, the speed of the particle is limited in the range of $[V_{\min}, V_{\max}]$ ($V_{\min} = -V_{\max} = -10$). The position should be in the allowable, and the optimal output \mathbf{gbest} is the global extreme.

The above basic PSO is simple while it is easy trapping into partial extreme and the accuracy of the searching is lower. The chaotic mapping iterative equation of Logistic used in this paper is:

$$\beta_j^{k+1} = \mu\beta_j^k(1 - \beta_j^k), k = 1, 2, \dots, \beta_j \in (0, 1), \beta_j \neq 0.25, 0.5, 0.75. \tag{10}$$

In the process of iteration, the optimal particle in every niche swarm is iterated chaotically by following latter equations:

$$\begin{cases} p_c = X_{j,\min} + \beta_j^{k+1}(X_{j,\max} - X_{j,\min}) \\ X_j^{k+1} = (1 - \lambda_m)X_j^k + \lambda_m p_c \end{cases}. \tag{11}$$

where, λ_m , called shrinkage factor, which determines the variation space of x_j , is from the following formula: $\lambda_m = 1 - [(m-1) / m]^u$, where u is the speed of shrinking ($u = 2$).

Now we can search for optimal multi-threshold by NCPSO. The concrete steps are as follows:

- a) Initialize the niche particle swarm. Randomly generate particles of number M and divide them into four son swarms.

- b) Calculate the fitness of every particle according to (6), find out the optimal particle of every niche particle swarm and the whole optimal particle;
- c) Calculate the distance between the optimal individuals $\mathbf{pbest}_i(m)$ and $\mathbf{pbest}_j(m)$ of the two particle swarm;
- d) If the iterative time meets the scheduled value, update and initialize the worst niche son swarm;
- e) Adopt chaotic mutation to the position of all niche optimal individuals to improve searching accuracy;
- f) Update the position and speed of every particle;
- g) If it meets the ending condition, stop the iteration, output the optimal multi-threshold and extract the image by this multi-threshold. If not, go to b).

4 Experimental Results and Analysis

A large number of experiments have been carried out to assess the performance of multilevel thresholding based on exponent gray entropy and NCPSO. We find that proposed method in this paper is superior in subjective visual by comparing with method in [13]. As the space is limited, one representative image is used to illustrate the conclusion.

At Fig. 1, (a) is the original image; (b)(c) are segment images processed by method in [13]; (d)(e) are segment images processed by proposed method. The optimal multi-threshold and running time are shown in Table 1, where the particle number of NCPSO, PSO both are 20, and the iterative times are 100. From the experimental results, we can see: (1) proposed method is significantly better than method in [13]. Being different from maximal Shannon entropy based only on histogram distribution, the exponent gray entropy reflects the uniformity of the gray scale immediately within the cluster. This makes the segment image more accurate. (2) from a subjective point of view on the visual effects, proposed method in this paper gets obvious advantages in details and texture of the segment images, such as levels between wall brick in the house image and the contour of peppers in the peppers image. (3) proposed method makes use of the swarm intelligence and search the optimal value through the competition and cooperation between particles, so it can avoid the prematurity, guarantee the accuracy of searching, and find the optimal multi-threshold by only few iterative time.

To further verify the effectiveness of the proposed method, we use an infrared thermal wave image (Composite propeller blade image) as the test image to illustrate. At Fig. 2, (a) is the original image; (b) is the segment image processed by three-threshold selection based on method in [13]; (c) is the segment image processed by three-threshold selection based on suggested method. And the optimal three-threshold and running time are shown in Table 2. We can see that, compared with the method of multi-threshold selection based on maximal Shannon entropy and PSO, segment image of the proposed method is more accurate in texture and clearer in details of feature. The experimental environment is Intel Celeration(R) CPU 2.66GHz/512MB memory/Matlab 7.0.

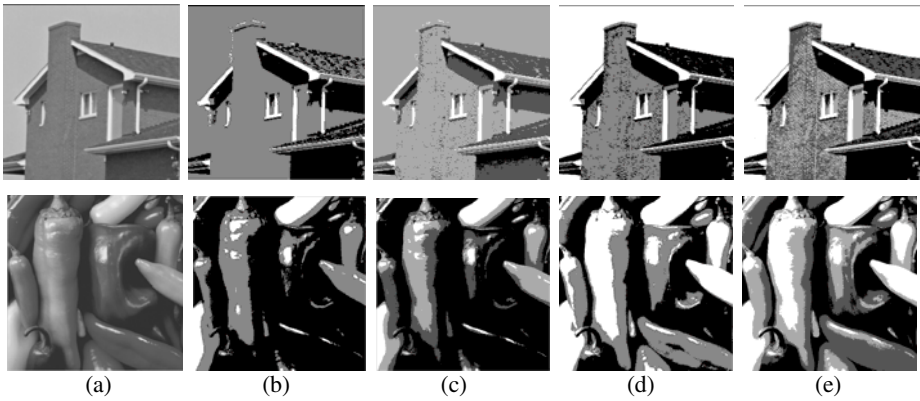


Fig. 1. House image(256×256) and Peppers image(512×512). (a) original images (b) two-threshold of method in [13] (c) three-threshold of method in [13] (d) two-threshold of proposed method (e) three-threshold of proposed method.

Table 1. The Optimal Multi-Threshold and Running Time

		Multi-Threshold / Running Time (s)	
		Two-threshold	Three-threshold
House	Method in [13]	(101,193)/0.594	(63,112,194)/0.656
	Proposed Method	(116,179)/0.500	(104,123,181)/0.578
Peppers	Method in [13]	(107,174)/0.625	(98,138,181)/0.687
	Proposed Method	(79,110)/0.515	(69,90,131)/0.609

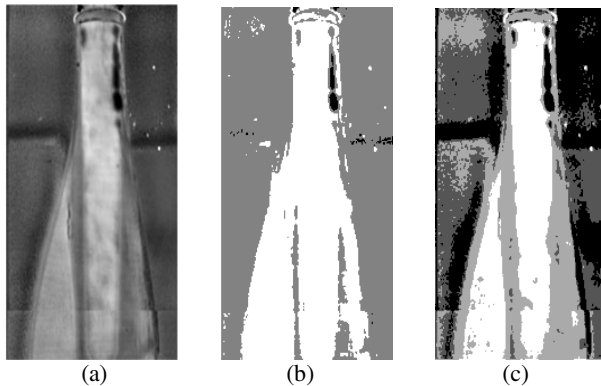


Fig. 2. Composite propeller blade image(299×155). (a) original image (b) three-threshold of method in [13] (c) three-threshold of proposed method.

Table 2. The Optimal Three-Threshold and Running Time

Composite propeller blade image	Three-threshold	Running Time (s)
Method in [13]	(33,101,166)	0.641
Proposed Method	(71,90,146)	0.565

5 Conclusion

Exponent gray entropy defined in this paper is different from maximal Shannon entropy or exponent entropy based only on histogram distribution. It reflects the uniformity of the gray scale directly within the cluster. Thresholding based on exponent gray entropy can improve the image segmentation results. NCP SO can not only guarantee the searching accuracy but also greatly reduce the cost of multi-threshold selection, and the operation speed is increased. Compared with the method of multilevel thresholding based on maximum Shannon entropy and PSO, suggested method has obvious advantages in the segmentation effect and running time.

Acknowledgments. Foundation: Open Foundation of the Key Laboratory of Nondestructive Testing (Nanchang Hangkong University), Ministry of Education (No.ZD2010290); National Natural Science Foundation of China (No.60872065).

References

1. Pal, N.R., Pal, S.K.: A Review on Image Segmentation Techniques. *Pattern Recognition* 26, 1277–1294 (1993)
2. Wu, Y.Q., Zhu, Z.D.: 30 Years (1962-1992) of the Developments in the Threshold Selection Methods in Image Processing (I). *Journal of Data Acquisition & Processing* 8, 193–201 (1993)
3. Wu, Y.Q., Zhu, Z.D.: 30 Years (1962-1992) of the Developments in the Threshold Selection Methods in Image Processing (II). *Journal of Data Acquisition & Processing* 8, 268–282 (1993)
4. Kapur, J.N., Sahoo, P.K., Wong, A.K.C.: A New Method for Gray-level Picture Thresholding Using the Entropy of The Histogram. *Computer Vision, Graphics and Image Processing* 29, 273–285 (1985)
5. Cao, L., Shi, Z.K.: An Automatic Multilevel Thresholding Method Based on Maximal Entropy. *Journal of Image and Graphics* 7, 461–465 (2002)
6. Cao, L., Shi, Z.K., Cheng, K.W.: Automatic Multilevel Thresholding Method Based on Maximum Entropy. *Transactions of Nanjing University of Aeronautics & Astronautics* 22, 335–338 (2005)
7. Cheng, H.D., Chen, J.R., Li, J.G.: Threshold Selection Based on Fuzzy C-partition Entropy Approach. *Pattern Recognition* 31, 857–870 (1998)
8. Luo, X.P., Tian, J.: The ICM Algorithm for Multi-level Threshold Selection by Maximal Entropy Criterion. *Journal of Software* 11, 379–385 (2000)
9. Chong, J.S., Zhou, X.K., Wang, H.Q.: Entropic Thresholding Method Based on Genetic Algorithm. *Journal of Beijing University of Aeronautics and Astronautics* 25, 747–750 (1999)

10. Sheng, G.F., Jiao, L.C.: Optimal Entropy Thresholding Image Segmentation Based on Genetic Algorithm. *Computer Engineering and Applications* 12, 103–105 (2003)
11. Wang, Y., Niu, Y.L., Tian, Y., Dong, J.Y., Hao, C.Y.: A More Stable and Accurate Genetic Algorithm for Segmentation of 3D Medical Images. *Journal of Northwestern Polytechnical University* 25, 442–445 (2007)
12. Wang, Y., Fan, Y.Y., Niu, Y.L., Monika, L., Qi, M., Hao, C.Y.: Effective Immune Genetic Algorithm for Segmentation of 3D Brain Images. *Journal of System Simulation* 20, 4136–4140 (2008)
13. Wei, M.M., Jiang, M.Y.: Multilevel Thresholding Methods for Image Segmentation Based on Particle Swarm Optimization. *Journal of Shandong University (Engineering Science)* 35, 118–121 (2005)
14. Xu, X.H., Zhang, A.: Entropic Thresholding Method Based on Particle Swarm Optimization for Image Segmentation. *Computer Engineering and Applications*, 8–11 (2006)
15. Ding, Y., Jin, W.Q., Zhou, H.F.: Multi-thresholds Image Segmentation Method Based on Particle Swarm Optimization. *Optical Technique* 34, 636–638 (2008)
16. Pal, N.R., Pal, S.K.: Entropic Thresholding. *Signal Processing* 16, 97–108 (1989)
17. Jia, D.L., Zhang, J.S.: Niche Particle Swarm Optimization Combined with Chaotic Mutation. *Control and Decision* 22, 117–120 (2007)

Learning Bag-of-Words Models Using Sparse Partial Least Squares

Jingneng Liu and Guihua Zeng

State Key Laboratory of Advanced Optical Communication Systems and Networks,
Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China
bz1jn@163.com, ghzeng@sjtu.edu.cn

Abstract. Representing images using Bag-of-Words (BOW) model has been shown excellent performance for image classification and retrieval. However, there are still some limitations in this model such as the presence of many noisy visual words and the hard to define vocabulary size. To circumvent these drawbacks, this paper concentrates on tuning compact, robust and thus efficient BOW model even with a universal size for image representation. The proposed approach increases expressive power by employing Sparse Partial Least Squares (SPLS) for tuning the traditional and high-dimensional BOW model and learning more discriminative subspace with 10 latent variables. The performance of learning BOW models to image classification is studied through extensive experiments on the VOC 2006 dataset. Empirical results indicate that the proposed method yields quite stable results, and outperforms the traditional BOW models with various vocabulary sizes and PCA with SVM.

Keywords: Bag-of-Words Model, Sparse Partial Least Square, Object Classification, Visual Vocabulary, Dimensional Reduction.

1 Introduction

Classifying images automatically into predefined semantic categories [1,2], i.e., classifying unlabeled images based on the presence or absence of instances of particular visual classes such as cars, people, bicycles, etc, has attracted increasing attentions in computer vision community, such as image classification, retrieval or annotation. However, automatic image categorization is still a challenge task, due to the appearance of foreground object instances varying substantially owing to changes in pose, viewpoint and scale, occlusions and within-class shape variations, imaging and lighting conditions, and background clutter. To overcome these problems, many attractive methods have been proposed. Nowadays, the state-of-the-art approach to these problems relies more and more on the concept of the Bag-of-Words model (BOW) applied to images. Despite its simplicity and lack of global geometry information, it still turns out to be surprisingly discriminatory, and has proven to be effective solutions for image object category classification [3] and video retrieval [4].

However, there are still some major limitations with the current Bag-of-Words models. The first one is the presence of many noisy visual words due to the coarseness of the unsupervised vocabulary construction process [5]. These vocabularies are

under- or over-representing the visual content, and may lead to high-degree multicollinearity. The second is the optimal parameter selection of the visual vocabulary size. The size of vocabulary is one of the most major parameters of the BOW model that is linked to the discriminative power of the model. On one hand, G. Csurka et al. [6] reported that performance improves steadily as the vocabulary grows. However, this can result in higher storage cost, time complexity and dimension disaster [7]. On the other hand, E. Nowak et al. [2] also show that there are initially substantial gains in performance as the vocabulary size increases, but overfitting becomes apparent for larger vocabularies. Similarly, J. C. van Gemert et al. [8,9,10] demonstrated that larger vocabulary sizes completely deteriorate the performance of the traditional BOW model. As the vocabulary size increases from 200 to over 80,000, Jun Yang et al. [11,12] experimentally revealed that the performance first rises dramatically, peaks at certain points, and after that either levels off or drops mildly. The other limitations are such that the loss of spatial geometric information arises in the classical Bag-of-Words model that seems prejudicial to the object category classification [5,13]. Furthermore, the popular construction process of the visual vocabulary is mostly unable to take object category information into account, and consequently the vocabulary may not be particularly informative to the object category classification task [14].

To overcome these limitations, several approaches are proposed by investigating the expressive power of vocabularies. J. Vogel and B. Schiele [15] described an image with a semantic vocabulary in meaningful visual words. Jurie and Triggs [1] obtained a proper vocabulary for an image using a data-driven approach characterized by unsupervised clustering, and specifically show that radius-based clustering outperforms the popular Kmeans clustering algorithm. Furthermore, Perronnin et al. [16] focused on concept-specific vocabularies, which endows each concept with its own unique vocabulary. Instead of hard-assignment, Jan C. van Gemert et al. [8,9,10] constructed compact vocabulary by using soft-assignment to model the uncertainty of the meaning of an image patch. Bosch et al. [17] conducted scene classification using a hybrid generative approach, and show that PLSA [18] improves performance upon LDA [19]. Similar to visual word co-occurrence modeling, extensive studies have demonstrated the effectiveness to overcome the limitation of BOW model with a large vocabulary size. For large-scale image retrieval systems, X. Zhang et al. [20] decomposed the BOW model of an image into three layers: background information, information related to a set of pre-learned topics and residual information. By packing Bag-of-Features to miniBoFs which makes use of Hamming Embedding (HE), H. Jégou et al. [21] represented per image only by taking 35KB of memory. The state-of-the-art system is newly reported by F. Perronnin et al. [22], which uses as little as a few hundreds of bits to represent per image by compressed Fisher vectors and significantly outperforms the recent retrieval technique using compressed BOV technique (miniBoFs). These previous works show that the expressive power and vocabulary size are two of the most important parameters of the popular BOW model, whereas a unified amount of vocabulary size for the same or different image collections and computer vision tasks can hardly be defined beforehand. This leads to necessary evaluations to determine the appropriate length of BOW model.

In this work, we concentrate on tuning compact and robust BOW model even with a universal size for image representation, and trying to retain or improve the ability to discriminate well between object categories. Inspired by the generative models such as PLSA and LDA based on classical BOW models, we propose a novel framework for achieving compact image representation, which learns discriminative BOW models based on traditional BOW models. Different from generative models by modeling visual word co-occurrence, the proposed method rebuilds a more powerful expressive BOW model by fairly removing the noise and multicollinearity among the initial visual words.

The nature of the proposed method makes an ideal setting for a statistical technique known as matrix factorization approaches, i.e., Partial Least Squares (PLS) [23,24,25]. However, to the best of our knowledge, there has not been any thorough evaluation of PLS for image classification. In this paper, we are interesting to introduce Sparse Partial Least Squares Regression (SPLS) [26,27] for learning compact BOW model by simultaneous dimension reduction and variable selection. The performance of learning BOW model to image classification is studied through extensive experiments on the VOC 2006 dataset. Empirical results indicate that learning BOW models with 10 latent variables yields quite stable results, and outperforms the classical BOW models with various vocabulary sizes and PCA with SVM.

The remainder of this article is organized as follows. In the next section, a brief overview of the SPLS algorithm is provided, and the proposed improvement method is described. In section 3, its impact on the object category classification performance is evaluated.

2 Proposed Method

2.1 Sparse Partial Least Squares for High Dimensional Data

The core of PLS [23] is a dimension reduction technique that operates under the assumption of a basic latent decomposition of the object category matrix ($Y \in R^{n \times q}$) and feature matrix ($X \in R^{n \times p}$). Let $Y_{n \times q}$ and $X_{n \times p}$ represent the column centered object category and feature matrices, respectively. PLS regression assumes latent components $T_{n \times k}$ underlying both Y and X. Hence, the PLS model is given by $Y = TQ^T + F$ and $X = TP^T + E$, where $P_{p \times k}$ and $Q_{q \times k}$ are coefficients (loadings) and $E_{n \times p}$ and $F_{n \times q}$ are errors. The latent components T are defined as $T = XW$, where $W_{p \times k}$ are K direction vectors ($1 \leq K \leq \min\{n, p\}$). The main machinery of PLS is to find these direction vectors. The k-th direction vector \hat{w}_k is obtained by solving the following optimization problem,

$$\max_w w^T M w \quad s.t. \quad w^T w = 1 \quad \text{and} \quad w^T S_{XX} \hat{w}_l = 0 \quad l = 1, 2, \dots, k-1, \quad (1)$$

where $M = X^T Y Y^T X$ and S_{XX} represents the sample covariance matrix of the predictors. For univariate PLS, this objective function can be interpreted as follows:

$$\max_w cor^2(Y, Xw) \var(Xw).$$

The SPLS [24, 25] incorporates variable selection into PLS by solving the following minimization problem instead of the original PLS formulation (1):

$$\min_{w,c} -kw^T M w + (1-k)(c-w)^T M(c-w) + \lambda_1 \|c\|_1 + \lambda_2 \|c\|_2, \quad (2)$$

subjected to $w^T w = 1$, where $M = X^T Y Y^T X$. This formulation promotes exact zero property by imposing L_1 penalty onto a surrogate of direction vector (c) instead of the original direction vector (w), while keeping w and c close to each other. Here, L_2 penalty takes care of the potential singularity of M . If the response Y is univariate, then the solution of this formulation results in a soft thresholded direction vector:

$$\hat{c} = (|Z| - \lambda_2/2)_+ \text{sign}(Z),$$

where $Z = X^T Y / \|X^T Y\|$ and $(x)_+ = \max(0, x)$. One can recast this soft thresholding as

$$\hat{c} = (|Z| - \eta \max_{1 \leq j \leq p} |Z_j|)_+ \text{sign}(Z),$$

where $0 \leq \eta \leq 1$. This η justifies the setting $0 \leq k \leq 0.5$ and $\lambda_2 = \infty$. Therefore, there are two key tuning parameters, η and K , in this formulation. Note that controlling η instead of the direction vector specific sparsity parameters λ_k , $k = 1, \dots, K$, avoids combinatorial tuning of the set of sparsity parameters and provides a bounded range for the sparsity parameter, i.e., $0 \leq \eta \leq 1$. Furthermore, if $\eta = 0$, then SPLS coincides with PLS.

Similarly to PCA and PLS, in the dimensionality reduction using SPLS, after relevant weight vectors are extracted, an appropriate classifier can be applied to the low-dimensional subspace. The difference between SPLS and PCA is that the former creates orthogonal weight vectors by maximizing the covariance between elements in X and y .

2.2 Learning Bag-of-Words Models

Fig. 1 shows the framework of the proposed learning BOW modes. The key idea behind applying BOW models for image representation is to quantize the continuous high-dimensional space of image features (e.g., SIFT descriptors) to a manageable vocabulary of visual words [14]. This is typically achieved by grouping the low-level features collected from a large image corpus into a specified number of clusters by using an unsupervised algorithm such as Kmeans, or radius-based clustering. By treating the center of each cluster as a visual word in a vocabulary, one can map each feature newly extracted from a real-world image onto its closest visual word, and represent the image as a feature vector by a histogram over the vocabulary of visual words, according to the presence or count of each visual word.

However, there are still some limitations in this classical BOW models, such as the presence of many noisy visual words and the hard to define vocabulary size. To circumvent these drawbacks, the proposed approach increases expressive power by employing SPLS for tuning the traditional and high-dimensional BOW model and learning more compact and discriminative subspace with 10 latent variables. In the learning stage, the SPLS automatically leads to the selection of relevant visual words, gives higher weights

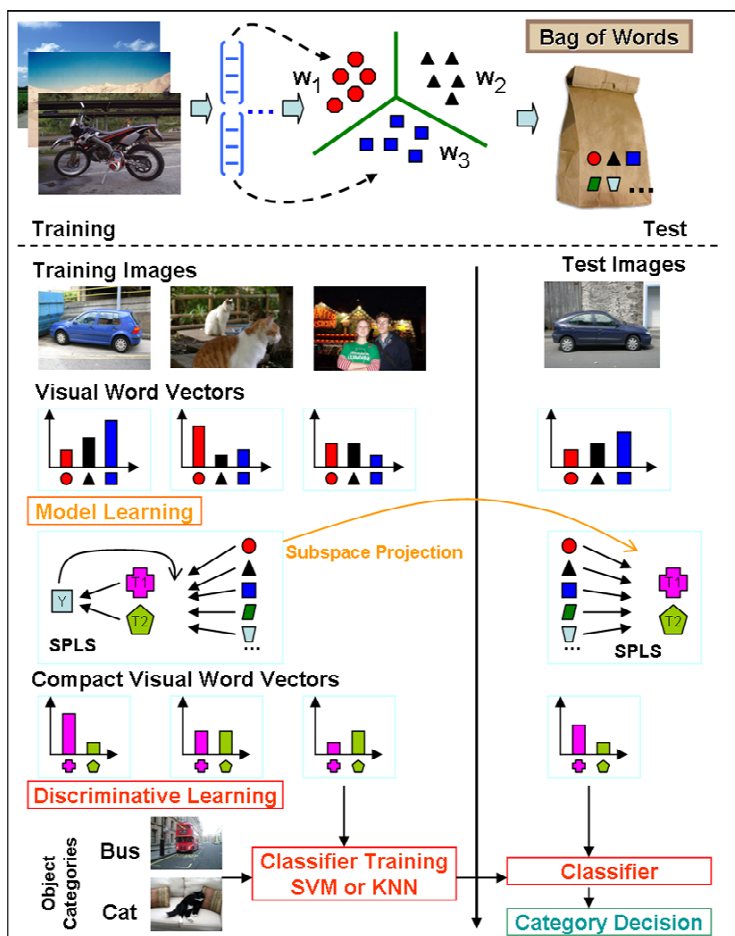


Fig. 1. Illustration of the proposed method employed for learning compact BOW model for image representation

to the visual words which contain more discriminatory characteristics, and filters out large numbers of noisy visual words. Therefore, this process can be seen as simultaneously dimension reduction and feature selection process depending on the visual words discriminatory, and can result in more compact and powerful BOW model for image representation.

Once the discriminative SMLS model has been estimated for the training images, the visual word vectors describing images are projected onto the weight vectors. The resulting low-dimensional features namely latent variables are the expected K components, and used to match a query sample during the testing stage. When a new sample is presented during the testing stage, its visual word vector is converted to latent variables by projecting onto the latent subspace that is previously estimated for the training images.

3 Experimental Evaluation

3.1 Dataset and Experimental Setup

To evaluate the performance of the proposed approach, the PASCAL Visual Object Classes Challenge 2006 protocol [28] is adopted. There are ten object classes: bicycle, bus, car, cat, cow, dog, horse, motorbike, person, and sheep. The classification task will be judged by the Receiver Operating Characteristic (ROC) curve. The used principal quantitative measure will be the area under curve (AUC).

The Hessian-Affine (HesAff) detector is used for extracting the keypoints and regions [29]. To capture important cues for object category classification, the low-level features for describing the keypoints and interesting regions employed only are the Scale-Invariant Feature Transform (SIFT) descriptor [30, 31]. In these experiments, the normalized region size is fixed to 41×41 pixels, and a 128D SIFT descriptor is computed. For region detection, normalization, and computing SIFT descriptors, the software routines provided by the matching protocol [32] are adopted.

3.2 Parameter Evaluation

Firstly, we investigate the parameter selection of SPLS, and Fig. 2 (a) illustrates its experimental results. In these experiments, the vocabulary size of the traditional BOW models for bicycle images is set to 1000, and the parameter K components of SPLS are set to a range of $\{4, 6, \dots, 44\}$, and the η is $\{0.1, 0.5, 0.9\}$. For Cross-Validation of SVM with RBF kernel, the parameter gamma is set to $\{-15, -12, -9, -6, -3, 0, 3\}$, and the cost is $\{-5, -2, 1, 4, 7, 10, 13\}$. The best AUC results obtained over SPLS subspaces is to use 10 for K components of SPLS, 0.5 for η , and 51 for k of KNN. However, the performance of the system drops when the number of latent variables increases beyond 10. This can be attributed to overfitting of the initial data of visual word vectors caused by using larger number of latent variables, and the KNN is essentially more suit for low-dimensional feature space. Secondly, we focus on the parameter selection of PCA,

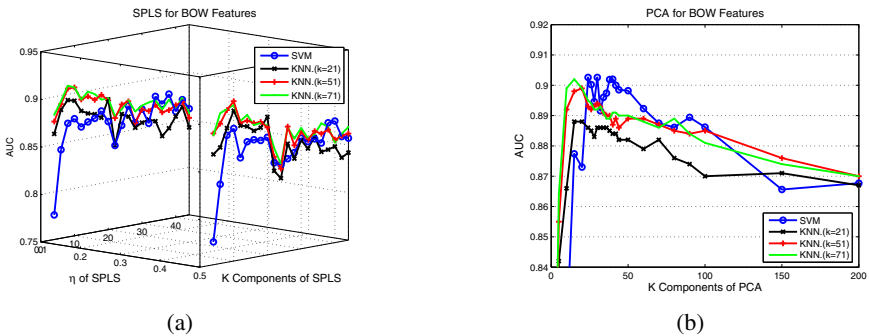


Fig. 2. Comparison of SPLS and PCA for dimensionality reduction. AUC curves for image classification according to the number of dimensions used to train the classifier. Training and test samples are bicycle images from VOC 2006 dataset.

and Fig. 2 (b) shows its experimental results. The dimensionality of PCA subspaces that always obtains better AUC results in other test cases is 44. Generally speaking, the results of KNN and SVM both peak and shock at a range about 20 ~ 44 of K components of PCA. After this range, the results obtained over PCA subspaces decrease slowly and gently.

3.3 Experimental Results

To gain insight in the performance variation between the various sizes of initial vocabularies and a fixed universal and compact subspace size tuned by PCA and SPLS, we experiment with initial vocabulary sizes varying with a range {200, 400, 500, 600, 1000} for PASCAL VOC 2006 dataset, 44 dimension for PCA subspace, and only 10 dimension for SPLS subspace. Three methods are compared, i.e., Kmeans + SVM, SPLS

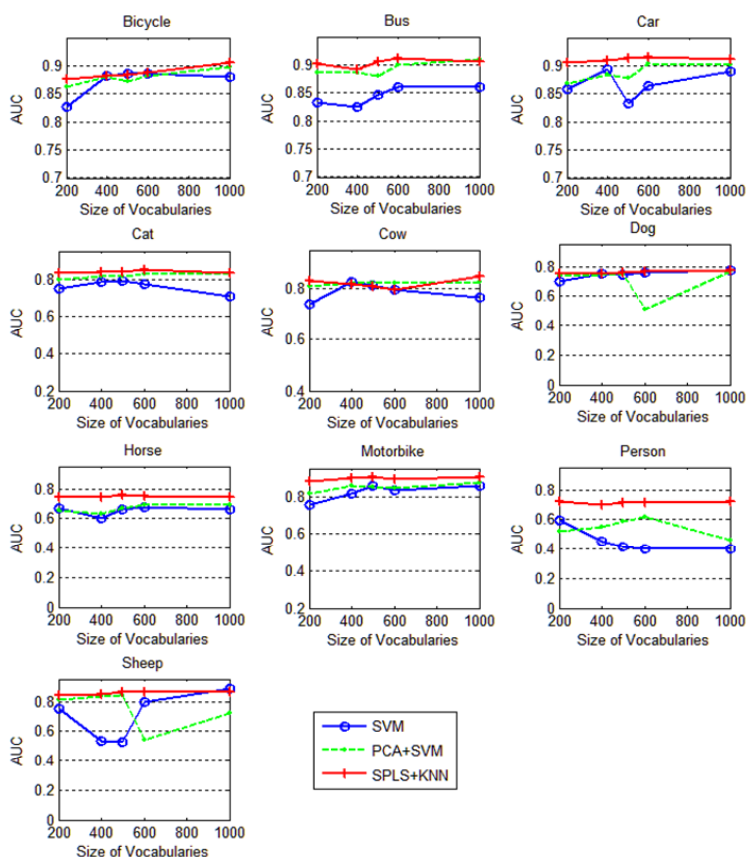


Fig. 3. Comparing the proposed learning BOW model with the traditional method. Performance evaluation results in AUC for the PCA and SPLS on PASCAL VOC 2006 dataset. Note that the dimensionalities of PCA and SPLS subspace are only 44 and 10 respectively, while the used traditional vocabulary sizes vary from 200 to 1000.

(10) + KNN.51 and PCA (44) + SVM. In order to make a clearly comparison among these methods between the initial various vocabulary sizes and the fixed subspace size, the AUC results of PCA and SPLS are plotted onto the same figures taken from Kmeans + SVM method, according to the subspace come from whom.

Fig. 3 shows the relationship between the classification performance and the visual vocabulary sizes or subspace size. It is clear that the vocabulary size has more or less impact on the classification performance. The first surprising observation we can make is that the proposed SPLS (10) + KNN.51 method clearly outperforms the Kmeans + SVM with various vocabulary sizes and PCA (44) + SVM over all of the test cases and all different vocabulary sizes. Another quite interesting observation is that the proposed SPLS (10) + KNN.51 method obtains fairly stable AUC results. However, the traditional Kmeans + SVM method shows unstable results over different vocabulary sizes, and a hypothesis is supported by the results in the bus, car and sheep test cases. Hence, learning BOW model by SPLS seems more robust to the variation of the initial vocabulary sizes. These notable facts illustrate that when more sophisticated tuning schemes for compact BOW model for image representation are employed, the impact of initial BOW model size turns to be not so significant. At this point, one can reach more robust and universal BOW model size for a special training set.

3.4 Computational Cost

Fig. 4 shows the comparisons of computational time among the SPLS, PCA and SVM methods for object category classification in bicycle test case. It is important to note that the time cost of SPLS (10) + KNN.51 seems insignificant in comparison to the other two methods. Thus, in addition to the superior performance, the low computational cost of subspace analysis makes SPLS more suitable for actual application than PCA and the traditional BOW models. Furthermore, another more dramatic observation is that the time cost of PCA (44) + SVM drops linearly, as the initial vocabulary size

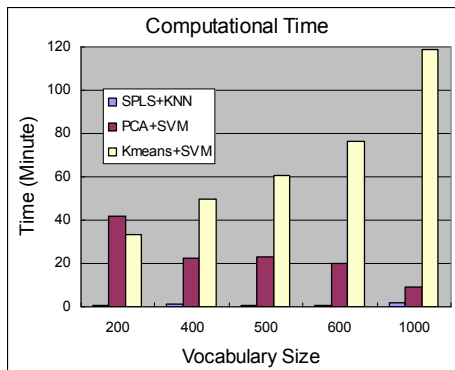


Fig. 4. Computational time comparisons considering the SPLS, PCA and SVM methods. The used traditional vocabulary sizes vary from 200 to 1000, and the dimensionalities of PCA and SPLS subspace are 44 and 10 respectively.

increases, though the dimensionality of the PCA subspace is still the same (44). The essence standing this observation is that the more expressive power the initial vocabulary exhibits, the more discriminative subspace the PCA achieves, and subsequently the less iteration counts and time cost the SVM requires. Hence, these results suggest that more expressive vocabulary subsequently benefits the PCA or SPLS to pursue more discriminative subspace.

4 Conclusions

This paper focuses on compact, robust, and thus efficient BOW model even with a universal size for image representation. Empirical results indicate that the proposed method rebuilds more compact BOW model, yields quite stable results, and outperforms the classical BOW models with various vocabulary sizes and PCA with SVM. While the traditional BOW model appears to be rather sensitive to the variety of the vocabulary sizes. Furthermore, low-dimensional subspace leads to both much lower computational and memory cost. Thus, in addition to the low computational cost, the superior performance makes SPLS more suitable for actual application than PCA and the traditional BOW models.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant No. 60773085, No. 60970109), and NSFC-KOSEF International Collaborative Research Funds (Grant No. 60811140346, and No. F01-2008-000-10021-0).

References

1. Jurie, F., Triggs, B.: Creating Efficient Vocabularies for Visual Recognition. In: Proc. 10th IEEE International Conference Computer Vision (2005)
2. Nowak, E., Jurie, F., Triggs, B.: Sampling Strategies for Bag-of-Features Image Classification. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part IV. LNCS, vol. 3954, pp. 490–503. Springer, Heidelberg (2006)
3. Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: A comprehensive study. *IJCV* 73(2), 213–238 (2007)
4. Sivic, J., Zisserman, A.: Video google: A text retrieval approach to object matching in videos. In: ICCV (2003)
5. Tirilly, P., Claveau, V., Gros, P.: Language modeling for bag-of-visual words image categorization. In: Proceedings of the 2008 International Conference on Content-Based Image and Video Retrieval, ACM CIVR 2008, Niagara Falls, Canada, pp. 249–258 (July 2008)
6. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: ECCV 2004 Workshop on Statistical Learning in Computer Vision, pp. 59–74 (2004)
7. Hastie, T., Tibshirani, R., Friedman, J.: *Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd edn., 745 pages in full color. Springer, New York (2009)

8. van Gemert, J.C., Geusebroek, J.-M., Veenman, C.J., Smeulders, A.W.M.: Kernel Codebooks for Scene Categorization. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part III. LNCS, vol. 5304, pp. 696–709. Springer, Heidelberg (2008)
9. van Gemert, J.C., Veenman, C.J., Smeulders, A.W.M., Geusebroek, J.M.: Comparing Compact Vocabularies for Visual Categorization. In: *Computer Vision and Image Understanding* (2010) (in press)
10. van Gemert, J.C., Veenman, C.J., Smeulders, A.W.M., Geusebroek, J.M.: Visual word ambiguity. *IEEE Trans. Pattern Analysis and Machine Intelligence* (2010) (in press)
11. Yang, J., Jiang, Y.G., Hauptmann, A.G., Ngo, C.W.: Evaluating bag-of-visual-words representations in scene classification. In: *Proceedings of the International Workshop on Multimedia Information Retrieval*, pp. 197–206. ACM, New York (2007)
12. Jiang, Y.-G., Ngo, C.-W., Yang, J.: Towards Optimal Bag-of-Features for Object Categorization and Semantic Video Retrieval. In: *ACM International Conference on Image and Video Retrieval (CIVR 2007)*, Amsterdam, Netherlands, July 9-11 (2007)
13. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: *IEEE CVPR* (2006)
14. Yang, L., Jin, R., Sukthakar, R., Jurie, F.: Unifying Discriminative Visual Codebook Generation with Classifier Training for Object Category Recognition. In: *Proceedings of Computer Vision and Pattern Recognition* (2008)
15. Vogel, J., Schiele, B.: Semantic modeling of natural scenes for content-based image retrieval. *International Journal of Computer Vision* (2010)
16. Perronnin, F., Dance, C., Csurka, G., Bressan, M.: Universal and adapted vocabularies for generic visual categorization. *IEEE Pattern Analysis and Machine Intelligence* (2010)
17. Bosch, A., Zisserman, A., Muñoz, X.: Scene classification using a hybrid generative/discriminative approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(4), 712–727 (2008)
18. Hofmann, T.: Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning* 42(1/2), 177–196 (2001)
19. Blei, D., Ng, A., Jordan, M.: Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, 993–916 (2003)
20. Zhang, X., Li, Z., Zhang, L., Ma, W.-Y., et al.: Efficient indexing for large scale visual search. In: *ICCV* (2009)
21. Jégou, H., Douze, M., Schmid, C.: Packing bag-of-features. In: *ICCV* (2009)
22. Perronnin, F., Liu, Y., Sanchez, J., Poirier, H.: Large-scale image retrieval with compressed Fisher vectors. In: *CVPR* (June 2010)
23. Wold, H.: Partial Least Squares. In: Kotz, S., Johnson, N. (eds.) *Encyclopedia of Statistical Sciences*, vol. 6, pp. 581–591. Wiley, New York (1985)
24. Schwartz, W.R., Kembhavi, A., Harwood, D., Davis, L.S.: Human Detection Using Partial Least Squares Analysis. Accepted to be presented in the *International Conference on Computer Vision (ICCV 2009)*, Kyoto, Japan, September 27-October 04 (2009)
25. Schwartz, W.R., Davis, L.S.: Learning Discriminative Appearance-Based Models Using Partial Least Squares. In: *Proceedings of the XXII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2009)*, Rio de Janeiro, Brazil, October 11-14 (2009)
26. Chung, D., Keles, S.: Sparse Partial Least Squares Classification for High Dimensional Data. *Statistical Applications in Genetics and Molecular Biology* 9, Article 17 (2010)

27. Chun, H., Keles, S.: Sparse partial least squares for simultaneous dimension reduction and variable selection. *Journal of the Royal Statistical Society: Series B* 72, 3–25 (2010)
28. Everingham, M., Zisserman, A., Williams, C., Gool, L.: The pascal visual object classes challenge 2006 results (2006), <http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2006/>
29. Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Van Gool, L.: A Comparison of Affine Region Detectors. *International Journal Computer Vision* 65(1/2), 43–72 (2005)
30. Mikolajczyk, K., Schmid, C.: A Performance Evaluation of Local Descriptors. *IEEE Trans. Pattern Analysis and Machine Intelligence* 27(10), 1,615–1,630 (2005)
31. Lowe, D.: Distinctive Image Features from Scale-Invariant Keypoints. *International Journal Computer Vision* 2(60), 91–110 (2004)
32. Mikolajczyk, K.: Affine Covariant Features, Visual Geometry Group, University of Oxford (2004), <http://www.robots.ox.ac.uk/~vgg/research/affine/>

Blind Watermark Algorithm Based on QR Barcode

Meifeng Gao and Bing Sun

Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education,
Jiangnan University, Wuxi 214122, China
mfkao2008@gmail.com, andysun2000@126.com

Abstract. In accordance with the structural features of QR barcode, a text blind watermarking algorithm is proposed in this paper. The algorithm is that the Hill₂ Encryption and BCH error correction encoding are used in the watermark pre-processing and then the watermark information is embedded in the rows and columns of QR barcode by using structural fine-tuning method. The QR barcode security and confidentiality are effectively improved by combining two-dimensional barcode with digital watermark technology.

Keywords: QR barcode, digital watermark, watermark encryption.

1 Introduction

Two-dimensional barcode, which is developed based on computer and information technology, is a standardized information storage and automatic identification technology. It is widely used because of its low cost, high reliability, large capacity and convenient to carry, etc.[1]. The two-dimensional barcode encryption technology is not a strict anti-counterfeiting technology, so digital watermark technology, which is used to improve the security performance of two-dimensional barcode, has become the research hotspot in recent years. The digital watermark combined with two-dimensional barcode is divided into two types: one is that the two-dimensional barcode information as a watermark is embedded into the cover image or text[2, 3], the other is that the digital watermark is embed into two-dimensional code[4-6]. In this paper, considering the structure of QR barcode (a matrix two-dimensional barcode), a blind watermark algorithm with the structural fine-tuning method to embed watermark in the rows and columns of the QR barcode is presented.

2 Watermark Pre-processing

Nowadays, the watermark types are generally text and image. Due to its low capacity, the text watermark has been widely used. In this paper, the text watermark is researched. Before the text watermark is embedded in QR barcode, it should be pre-processed. The watermark pre-processing includes the following steps: ① In order to enhance the security of watermark, Hill₂ Encryption[7], which can achieve the same letter in different locations of a text correspond to different cipher text letter using a

reversible transformation matrix, is used to encrypt the text. ② The information is compressed to reduce the data capacity according to the data redundancy. ③ The information is error correction encoded by using BCH (15,5) code[8], so that it has automatic error correction function.

3 Watermark Embedding Algorithm

For the purpose of increasing security of two-dimensional barcode, digital watermark, such as: biological characteristics, image, text, etc., can be embedded into the two-dimensional barcode. The algorithm, in which the digital watermark is embedded into QR barcode, is as follows.

3.1 Barcode Pre-processing

Fig.1 is a QR barcode pre-processing schematic diagram.

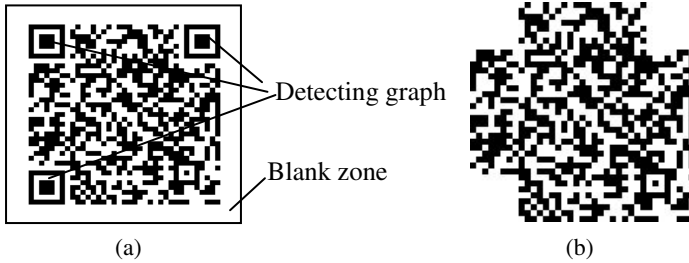


Fig. 1. QR barcode pre-processing schematic diagram: (a) Standard QR barcode, (b) Removed the blank zone and detecting graph

In the blank zone around a QR barcode, any pixel change will make a serious distortion of the QR barcode image, thus the blank zone must be removed. In addition, the detecting graph of QR barcode belongs to visually sensitive area, which occupies a large area and has invariance, small change in it is easily detected, so the detecting graph also need to be removed. Fig.1(b) is an effect diagram of the blank zone and detecting graph removed.

3.2 Embedding Capacity

Known current block is module A, black pixel value is 0, and white pixel value is 1. The $ave(A)$ is the average pixel value of module A, the width of module A is set to Q . The module diagram is shown as Fig. 2.

A1	A2	A3
A4	A	A5
A6	A7	A8

Fig. 2. The module diagram

According to the structure characteristics of QR barcode, the watermark is embedded into the rows and columns of QR barcode. Considering the robustness and invisibility, the six embedding conditions are defined.

① Row embedding conditions

Case 1: $\text{ave}(A) < 0.5$; $\text{ave}(A5) > 0.5$; $\text{ave}(A2) > 0.5$; $\text{ave}(A7) > 0.5$.

Case 2: $\text{ave}(A) < 0.5$; $\text{ave}(A5) > 0.5$; $\text{ave}(A2) > 0.5$; $\text{ave}(A8) < 0.5$.

Case 3: $\text{ave}(A) < 0.5$; $\text{ave}(A5) > 0.5$; $\text{ave}(A7) > 0.5$; $\text{ave}(A3) < 0.5$.

② Column embedding conditions

Case 4: $\text{ave}(A) < 0.5$; $\text{ave}(A7) > 0.5$; $\text{ave}(A4) > 0.5$; $\text{ave}(A5) > 0.5$.

Case 5: $\text{ave}(A) < 0.5$; $\text{ave}(A7) > 0.5$; $\text{ave}(A4) > 0.5$; $\text{ave}(A8) < 0.5$.

Case 6: $\text{ave}(A) < 0.5$; $\text{ave}(A7) > 0.5$; $\text{ave}(A5) > 0.5$; $\text{ave}(A6) < 0.5$.

The realization of various situations are shown in Fig.3 and Fig.4, symbol X represents any color. The row embedding capacity is recorded as row_total, and column embedding capacity is recorded as col_total.

A2	A3	white	X	white	X	X	black
A	A5	black	white	black	white	black	white
A7	A8	white	X	X	black	white	X

(a) (b) (c) (d)

Fig. 3. Row embedding positions: (a) embedding module, (b) Case 1, (c) Case 2, (d) Case 3

A4	A	A5	white	black	White
A6	A7	A8	X	white	X

(a) (b)

White	black	X	X	black	white
X	white	black	black	white	X

(c) (d)

Fig. 4. Column embedding positions: (a) embedding module, (b) Case 4, (c) Case 5, (d) Case 6

3.3 Watermark Pre-processing

Before the watermark is embedded in QR barcode, the input text is encrypted by Hill₂, compressed, and error correction encoded by BCH(15,5). Assuming the input text length is L. The pre-processed watermark text length is 15L. In order to make the watermark be embedded normally, it is ensure that the input text length L satisfies (1).

$$15L \leq \text{row_total} + \text{col_total} . \quad (1)$$

3.4 Watermark Embedding

Testing each module of QR barcode, if the module satisfies the embedding conditions, the watermark is embedded in it. For row embedding, if embedding 1, a Q*W black pixels block is increased in the module A and module A5 border, here Q is the width of module A, the size of W can be modified according to the actual situation. If embedding 0 without any change. Similarly, for column embedding, if embedding 1, a W*Q black pixels block is increased in A and A7 boundary, otherwise unchanged.

4 Watermark Extracting Algorithm

Because the watermark is extracted without the original QR barcode image, the watermark extracting is blind extraction. The QR barcode image embedded watermark is easy to geometric distortion and tilt when it is collected or printed. In this paper, the bilinear interpolation transform[9] is used to recover the distorted barcode, the binary morphological erosion and dilation are used to repair the holes and cracks generated during forming the corrected QR barcode. When the corrected OR barcode image is got, the watermark can be extracted from it.

Fig.5(c) shows the difference between the standard barcode contained watermark and the shooting image after bilinear correction. The white parts indicating the pixel value not the same shows that some of the pixel value shift occurred in the corrected QR barcode image. So it is not feasible to extract watermark through the reverse steps of embedding algorithm. From Fig.5(d) we can found that the watermarked module length is slightly longer than the normal module. The useful watermark extracting algorithm is as follows.

Firstly, according to the module width N and embedding conditions, all the embedded points are found. On the assumption that the module A is likely to row embedded point, setting counter Num = 0, and to the center column of module A as base, calculating the column mean to left and to right respectively. If the column mean is less than 0.5, then Num++, until the column mean is greater than 0.5. If Num is satisfies (2), it denotes that embedded data is 1, otherwise is 0.

$$\text{Mod} (\text{Num}, k) < T . \quad (2)$$

Where, $k = N + T/2$, N is module width of the corrected QR barcode. $T = N*W/Q$, is threshold, which is the watermark width of the corrected QR barcode.

Similarly, if module A is likely to column embedded point, setting the center row of module A as base, and calculating the row mean to up and to down separately.

After the watermark information is extracted, BCH(15,5) error correction decoding, compression decoding and Hill2 decryption are used to get original text information.

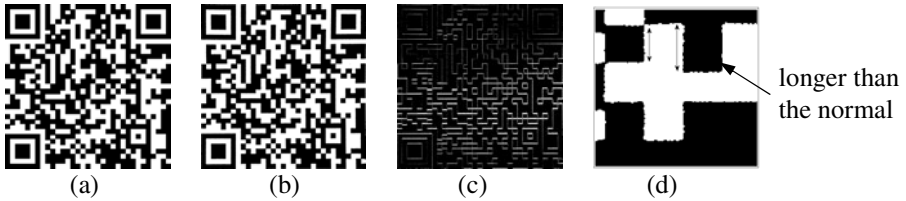


Fig. 5. Comparison between standard barcode and correct barcode: (a) Standard barcode contained watermark, (b) Correct barcode, (c) Difference image, (d) Details of the barcode module

5 Experiments

Firstly, the standard QR barcode image is generated, barcode version is 4, error correction level is 25%, the module size Q is set 4, encoding information is “jiangnan university is a good school”. The embedding capacity of QR barcode is calculated according to the embedding conditions, $row_total=119$, $col_total=106$, So the maximum length of the embedded text is $L=15$ according to formula(1). The original watermark text is “jnrain” and “watermarks” separately. In the Hill₂ encryption period, the key matrix $\mathbf{K} = [1, 2; 0, 3]$. The watermark width is set as $W = 1$. In order to verify the effectiveness and robustness of the algorithm, two attack types: staining and shooting, are employed. The experiment shows that watermark information, which are “jnrain” in Fig.6(a) and “watermarks” in Fig.6(b), can be directly extracted with the watermark extraction algorithm for staining attack.

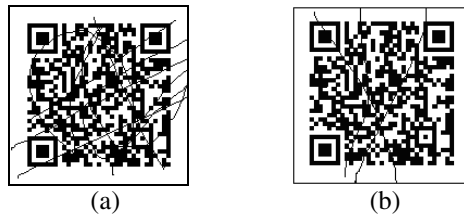


Fig. 6. Staining attack: (a) with “jnrain” watermark, (b) with “watermarks” watermark

But for shooting attack, the bilinear interpolation transform is needed. Fig.7 is the specific transformation process. The module width of the corrected QR barcode, being shown in Fig.7(c), is 9. The experiment shows that the watermark “jnrain” can be extracted and decoding correctly by the blind watermark extracting algorithm after the bilinear interpolation transform. In addition, the QR barcode embedded watermark also can be read correctly.

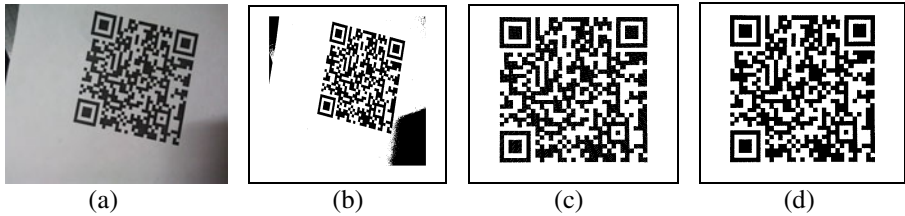


Fig. 7. QR barcode with watermark pre-processing: (a) Shooting attack sample. (b) Binarization, (c) Bilinear transformation, (d) Morphological repair

6 Conclusion

With the rapid development of information technology, two-dimensional barcode is widely used. To improve QR barcode security, the Hill₂ is used to pre-encrypt watermark text, and BCH(15,5) error correction coding is used for the watermark robustness. The experiments show that the algorithm has high robustness for staining and shooting attack, and is also simple and low complexity. This technique can be used to identify the authenticity of two dimensional barcode, and improve the safety and suitability of two-dimensional barcode.

References

1. Zhang, H., Zhang, D., Zhao, S.: Bar Code Technology and Application. Tsinghua university press, Beijing (2009)
2. Chung, C.-H., Chen, W.-Y., Tu, C.-M.: Image Hidden Technique Using QR-Barcode. In: 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 522–525. IEEE Computer Society, Piscataway (2009)
3. Chen, W.-Y., Wang, J.-W.: Nested image steganography scheme using QR-barcode technique. *Optical Engineering*, SPIE, 057004-1–057004-10 (2009)
4. Chen, Z., Yao, Y.-H., Wang, X.J.: Digital watermarking technique based on 2-D Barcode. *Computer Applications*, 1998–2000 (2006)
5. Bo, Z., Jin, H.: Information hiding technique based on PDF417 barcode. *Computer Engineering and Design*, 4806–4809 (2007)
6. Chao, Y.-H., Liu, L.-S., Xue, L.-Q., et al.: Information Hiding Algorithm Based on PDF417 Barcode. *Computer Engineering*, 131–133 (2010)
7. Ismail, I.A., Mohammed, A., Hossam, D.: How to repair the Hill cipher. *Journal of Zhejiang University SCIENCE A*, 2022–2030 (2006)
8. Wang, L.-X., Guo, B.-Z.: A New Encoding and Decoding Scheme of BCH Code. *Journal of Hebei university (Natural Science Edition)*, 161–164, 207 (2002)
9. Yang, J.-L., Gao, M.-F.: Research on Binarization Algorithm for QR Code. *Computer Engineering and Applications*, 176–178 (2009)

Cloth Pattern Simulation Based on a $1/f$ Noise Method

Beibei Li¹ and Zhihong Zhao²

¹ School of Electronic & Electrical Engineering, Shanghai Second Polytechnic University,
Shanghai 201209, China
Libeibei2@163.com

² School of Energy & Electronic Engineering, Nanjing Institute of Industry Technology,
Nanjing, 210046, China
zhaozhihong595@163.com

Abstract. This study addresses a cloth pattern simulation method on the basis of a $1/f$ noise, which is used to generate 3D controlled curves for the objective cloth. Users can conveniently get the different cloth pattern effects only by controlling the key points' positions of the $1/f$ noise function. The curved surface can be fitted by these 3D controlled curves. For the purpose of enhancing the visualization, a texture mapping method is employed and some improvements are made upon the illumination model. Finally, the experimental results show that the method is efficient and has potential research importance in the cloth texture simulation.

Keywords: cloth, pattern, $1/f$ noise.

1 Introduction

It has been an increasingly popular issue for decades to research on addressing new theories and technologies to simulate or animate the texture pattern of one or more cloth materials. In computer vision and computer graphic fields, the application of cloth pattern simulation has yet to be fully explored. Currently, simulation for cloth pattern follows the two types of trend. The first is to simulate the 3D illumination on the basis of the 2D illumination model, e.g., Refs. [1,2]. The second is to attend to simulate the cloth pattern on the basis of its 3D model, e.g., Ref. [3, 4]. YAN Gang Feng and YANG Bao Zhu address one improved Phong algorithm, which may be realized and processed in computer. On the 2D illumination model, they simulated its 3D illumination effect [5]. Unfortunately, the focus of their algorithm is only to simulate the cloth weave structure. They fail to realize its geometrical model simulation. MA Ling Zhou, XU Duan Qin and CHEN Chun present a method for cloth simulated in 2D virtual environment based oil mesh model [6]. Their algorithm follows two steps, one of which is to establish the model of pseudo-curving surface meshes and the second is to map the pattern of the cloth to the pseudo-curving surface. The simulation may meet the user's requirement. But they can not properly represent the cloth wrinkles. In addition, their pattern in the virtual environment should still be single. SHEN Jian bing, JIN Xiao gang, MAO Xiao yang and FENG Jie qing propose an

approach for designing a variety of textures from a single sample texture [7]. In their method, the original small texture is segmented into some layers. Each layer contains one special texture element, and it may be deformed by a set of chaotic-based transformation per formations. These deformed layers can be added together to shape a new texture. They adopt the graph cut algorithm and some image inpainting technique to patch the overlapping regions. In addition, by using the optimized graphcut synthesis algorithm, a variety of textures from a single texture may be generated. Unfortunately, they cannot design video textures through deformation.

In this work, we discuss cloth pattern simulations based upon a $1/f$ noise method. The rest of this paper is organized as follows. In section 2, we give a 3D cloth model built by the controlled curves. In the sub-sections of 2.1 and 2.2, we address the 2D curve fitting and the corresponding 3D surface fitting operations. Then we address the texture mapping in order to enhance the visual effect in section 3. In the section 4, we discuss the illumination model and the corresponding experimental results. Finally, we summarize the conclusions in section 5.

2 Three-Dimensional Model

The cloth 3D model is built by the controlled curves, which may be generated by fitting such series of key points. We use Perlin noise to simulate the cloth controlled curves. Perlin noise is usually a stochastic function, used to generate some sine functions with continuously changing amplitudes. It is employed to simulate the nature dynamics. The basis geometric model may be shown in Figure.1.

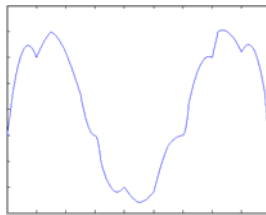


Fig. 1. The basis geometric model

2.1 Two-Dimensional Curve Fitting

In the $1/f^\beta$ noise, we get $1/f$ noise when $\beta=1$. In time-domain, $1/f$ model is based on sine function, in which its amplitude will correspondingly change when its phase changes. For the purpose of approximately representing the cloth properties, piecewise linear functions are employed to define the relation between the amplitude and the phase. We assume that the amplitude is represented by A , the phase represented by x , $x \in [300, 1200]$. The objective curve may be controlled by such key points. Here, we use 19 key points to express the coarse curve shape. For instance, the coordinates of the first curve's key points can be given in Table 1.

Table 1. The coordinates of the first curve's key points

Key points	(x_i, y_i)
(x_0, y_0)	(300, 0)
(x_1, y_1)	(320, 10)
(x_2, y_2)	(400, 15)
(x_3, y_3)	(450, 20)
(x_4, y_4)	(500, 16)
(x_5, y_5)	(550, 8)
(x_6, y_6)	(600, 0)
(x_7, y_7)	(620, -5)
(x_8, y_8)	(700, -10)
(x_9, y_9)	(750, -13)
(x_{10}, y_{10})	(800, -11)
(x_{11}, y_{11})	(850, -2)
(x_{12}, y_{12})	(900, 0)
(x_{13}, y_{13})	(920, 6)
(x_{14}, y_{14})	(1000, 15)
(x_{15}, y_{15})	(1020, 20)
(x_{16}, y_{16})	(1100, 16)
(x_{17}, y_{17})	(1180, 8)
(x_{18}, y_{18})	(1200, 0)

Now, we assume that A_i ($i=1, 2, 3, \dots, 18$) of the 1/f noise linear change linearly, expressed as

$$A = k_i x + b_i. \quad (1)$$

Where, x is the x -coordinate of the corresponding point; k_i, b_i are the parameters of the corresponding parts, defined by the coordinates of their start and end points. According to y variations, we define the curve phase Φ varies within the range of $0 \sim 1.5\pi$. Thus, the relation between x and Φ may be expressed as

$$\Phi = 1.5\pi(x - x_0) / (x_{18} - x_0). \quad (2)$$

Where x_0 and x_{18} are the x coordinates of the start and end points. Based on the y coordinates of the 19 controlled points, the corresponding A values are determined by

$$A = y \sin^{-1} \Phi. \quad (3)$$

Therefore, according to its controlled points of $(x_{i-1}, y_{i-1}), (x_i, y_i)$ and the amplitudes of A_{i-1}, A_i , we may find out the linear relation of the i -th among all of the 18 curves. Correspondingly, k_i and b_i in Eq.1 can be derived from

$$A_n = y_n \sin^{-1}(1.5\pi(x_n - 300)/900). \quad (4)$$

Where, $n=x_{i-1}, x_i$. Then we get

$$k_i = (A_i - A_{i-1}) / (x_i - x_{i-1}), \text{ and} \quad (5)$$

$$b_i = A_{i-1} - k_i x_{i-1}$$

2.2 Three-Dimensional Surface Fitting

Assume that the m curves $L_1, L_2, L_3, \dots, L_{m-1}, L_m$ have been fitted by the $1/f$ noise model in Section 2.1, based on which we may build its 3D model. The z coordinates of such m curves are ascending sorted as $z_1, z_2, z_3, \dots, z_m$. For the purpose of enhancing the smooth visual effects, we employ an interpolation method to interpolate n curves between the two curves. Now, we totally get $(m-1)n$ curves. We may give the definition of the relation between the key points on these n curves according to the relation between the coordinates of the corresponding controlled points on the two curves.

Here, we give a detailed description of the smooth processing between the first and the second curves. Among the curves $L_1, L_2, L_3, \dots, L_{m-1}, L_m$, we assume P_{2-1} is the first key point of L_2 . Correspondingly, P_{1-1} is that of L_1 . Furthermore, the x coordinate values of P_{1-1} and P_{2-1} are exactly the same. The x coordinates ranges of the key points on L_2 are k_x times of those on L_1 and meanwhile the y coordinate values of the controlled points on L_2 are k_y times of those on L_1 . Between L_1 and L_2 , the new interpolated curves are numbered as $l_{1-1-2}, l_{1-2-2}, l_{1-3-2}, \dots, l_{1-(n-1)-2}, l_{1-n-2}$. All of these interpolated curves' x coordinates initial values are constant. Therefore, the key points' x coordinate range of l_{1-i-2} is $k_x i/(n+1)$ times of that of L_1 and meanwhile its y coordinate value is $k_y i/(n+1)$ times. Based on the method mentioned above, the cloth can be modeled, shown in Fig.2.

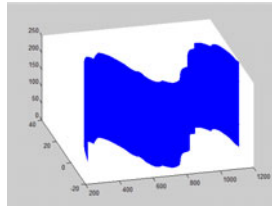


Fig. 2. The 3D modeled cloth

In Fig.2, we select six curves as the base curves. We interpolate 50 new curves between these six curves. On each new curve, we select totally 19 controlled points in order to define their coarse shapes. The phase values of these correspondingly controlled points are given the same values. This will yield to the small deformations while the stochastic property is not still satisfied. In actual process, the key points' positions and the number can be properly adjusted according to the cloth physical model, material properties, and so on. When all the work is done, the cloth properties may be properly represented and the better results have been achieved.

3 Texture Mapping

Texture mapping is a significant process to allow you to use a variety of tools to create visual-interesting patterns to be displayed on the simulated object. In texture mapping operations, there is a key point to be deeply mastered that is how to deal

with the model space and the objective space. In this study, we employ texture mapping to realize the 2D texture pattern mapped on to the 3D model. Firstly, we may load images texture pattern using *Imread* command, and then employ warp command to realize the texture pattern mapped onto the model. The algorithm is not complex and therefore should not be discussed here in detail.

4 Illumination Model

On the basis of the parametric surface, we make some improvements on the 3D illumination model so that the cloth plane illumination may be animated. Finally, the simulation is made by combining the texture mapping method for the 2D and 3D illumination models. In 3D computer graphic field, Schlick's BRDF model is a formula for approximating the BRDF model of metallic surfaces. In this study, we adopt a Schlick's BRDF illumination model, shown as

$$f = (1-G(u)G(v)) \pi^{-1} + 0.25G(u)G(v) \pi^{-2} u^{-1} \varrho^{-1} (1 + \varrho t^2 - t^2) \quad (6)$$

and

$$G(x) = x(\varrho - \varrho x + x)^{-1} \quad (7)$$

Where, t , u and v are the corresponding cosine values of the three angles α , θ and θ' , respectively; $G(u)$, $G(v)$ may be defined by Eq.(7); f is the reflection light intensity received by the observer; ϱ is the coarseness of the cloth surface, and its range varies within $[0,1]$. If $\varrho=0$, the object's surface is smooth and only specular reflection happens without any diffuse reflection. Otherwise, there is only diffuse reflection without any specular reflection when $\varrho=1$. The method has been implemented and tested for 2D and 3D models under MATLAB R2010b software background on a PC with a 2.60 GHz Pentium (R) Dual-Core CPU and 2.00 GB memory. In our implementations, we mainly use Curve Fitting Toolbox and Image Processing toolbox in MATLAB software. One of the experimental results may be shown in Fig.3.

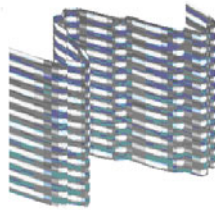


Fig. 3. Cloth pattern simulation result

5 Conclusions

This study proposes a 3D cloth geometric model on the basis of a $1/f$ noise mathematical model. On the curve, we select the properly representative key points according to the cloth special properties. The curve may be fitted on the basis of $1/f$ noise. We

use sine functions to model the curve, and their amplitudes are linearly controlled by the key points' geometrical positions. When the curved surface is fitted, we can control the amplitudes depending upon the key points' positions. Then, the texture pattern can be mapped onto the curved surface. In order to improve the visualism, the Schlick's BRDF model is employed in the study. Users can conveniently get the different cloth pattern effects only by controlling the key points' positions of the $1/f$ noise function. The comparison with the other similar researches shows that the compiled program of this technique is easier to realize and have a good flexibility.

Acknowledgement. It is Supported by Leading Academic Discipline Project of Municipal Education Commission, Project Number : J51801.

References

1. Xu, Y.Q., Liu, S.Q., Qi, D.X.: Generation of Fabric Textures. *Journal of Software* (6), 409–412 (1998)
2. Ma, Y.F., Wang, X.H.: Fabric Texture Feature Extraction Based on Wavelet Theory. *Journal of Hangzhou Institute of Electronic Engineering* (6), 85–88 (2002)
3. Lin, W.-C., Liu, Y.: Tracking Dynamic Near-Regular Texture Under Occlusion and Rapid Movements. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006, Part II*. LNCS, vol. 3952, pp. 44–55. Springer, Heidelberg (2006)
4. Liu, H., Chen, C., Shi, B.-L.: Simulation of 3D Garment Based on Improved Spring-Mass Model. *Journal of Software* 14(3), 619–627 (2002)
5. Yan, G.F., Yang, B.Z.: Light Model of Fabrics and Its Computer Simulation. *Journal of Zhejiang Institute of Science and Technology* 16(3), 198–203 (1999)
6. Ma, L.Z., Xu, D.Q., Chen, C.: Cloth Simulated in 2D Virtual Environment Based on Mesh Model. *Journal of Image and Graphics* 8(8), 964–969 (2003)
7. Shen, J., Jin, X., Mao, X., Feng, J.: Completion-based texture design using deformation". *Visual Computer* 22, 936–945 (2006)

An Anti-noise Determination on Fractal Dimension for Digital Images

Ying Shi, Shu Cheng, Shuhai Quan, and Ting Bai

School of Automation, Wuhan University of Technology, China
{a_laly, cheng1988shu}@163.com

Abstract. Since most images coming from nature show the fractal characteristic, the fractal dimension can be used to quantitative characterization and analysis on these images. However, noise may have an effect on the fractal dimension. In this work, the effects of salt & pepper noise, Gaussian white noise and multiplicative noise on the fractal dimension are studied by using different material images. The study shows that the three kinds of noises all have a significant effect on the fractal dimension. Digital images with added noise become coarser, and their corresponding fractal dimensions increase. An anti-noise determination on fractal dimension based on differential box counting (DBC) algorithm and noise character is proposed. Pixel values in a box are all used and the deviation between the maximum and minimum within the box is replaced with double standard deviation. Compared with the general pretreatment method, the proposed method is valid and convenient.

Keywords: fractal dimension, anti-noise, differential box counting algorithm.

1 Introduction

Fractal theory has gradually become a critical subject since its inception in 1970, and is widely used in physics, meteorology, geomorphology, materials science, social science, art and so on. Recently, many researchers have devoted themselves to study fractal characterization for different materials [1-4]. Most nature gray images have fractal characters [5], and the most important parameter of fractal is fractal dimension, which is strictly defined by F. Hausdorff and is not a whole number. The fractal dimension is useful in quantificational characterization of the object feature, and plays a critical role in texture analysis, image segmentation, and shape description and so on. Thus, accurate calculation on fractal dimension is very important in these applications.

At present, the fractal dimension is obtained by taking some experiment on fractal graphics or fractal phenomena. Due to different objects and methods, the commonly-used measurements on the fractal dimension are taken by using scaling, correlation function, distribution function, spectrum and different observed scale [6]. Box-dimension approach and differential box-counting approach are both realized by using different observed scale. Unfortunately, it should be noted that there are some issues for the method based on experimental data. At first, the obtained experiment data is usually too few to calculate an accurate statistic, fractal dimension. Secondly, the

method based on experiment data is not suitable for quantificational characterization on object feature because in this way fractal dimension does not have a clear physical meaning. Consequently, the determination method based on image is the most common one. Thereinto, the differential box-counting (DBC) algorithm via changing observed scale is proved to be the least computational complexity, and it gives an easy way to compute fractal dimension [7]. In general, during collection and communication images are inevitably disturbed by noises, which will affect the fractal dimensions of images and can not be ignored. The aim of this paper is to provide a modified algorithm to deal with such problem. We will study the relationship between different noises and fractal dimension, and then propose a robust determination algorithm of differential box-counting dimension.

2 The Principle of DBC Algorithm and Its Anti-noise Performance

2.1 The Principle of DBC Algorithm

DBC algorithm was initially proposed by Sarkar and Chaudhuri [7], and its application in images is as follow. Consider the image of $M \times M$ pixels as a three-dimensional (3D) space with (x,y) indicating the two-dimensional (2D) position and the third coordinate (z) denoting gray-level [11]. Hence, the (x,y) space is divided into grids of size $S \times S$ and the third dimension (z) is the small pieces with length S' , thus, each grid comprise a column of boxes of $S \times S \times S'$. S' should meet $[G/S'] = [M/S] = 1/r$, where G is the total number of gray-levels. Assume that the maximum and the minimum gray-level of the image in the $(i,j)^{th}$ grid fall in the box number l and k respectively. Then, the $(i,j)^{th}$ grid distribution n_r with size r is counted by $n_r(i,j)=l-k+1$, and the total number of boxes N_r is simply the summation of n_r from all grids, definition is as follow,

$$N_r = \sum_{i,j} n_r(i, j) \quad (1)$$

N_r values are different corresponding to different r , the fractal dimension D is estimated using Eq. (3) from least square linear fitting of $\log(N_r)$ and $\log(1/r)$.

$$D = \log N_r / \log(1/r) \quad (2)$$

2.2 The Effect of Noise on the DBC Algorithm

Computed tomography (CT) images of glass fiber and scanning electron microscope (SEM) images of TGP-H-060 carbon fiber paper and asphalt are obtained for this study, and they all have better fractal characteristics. First, convert these images into BMP images. Then, add salt and pepper noises with different densities, Gaussian random noises and multiplicative noises with zero mean and different variances to these images for the simulation of common grain noise and electronic noise. Last, calculate their fractal dimensions by using DBC algorithm, and the computation results of images without and with noise are shown in Fig. 1.

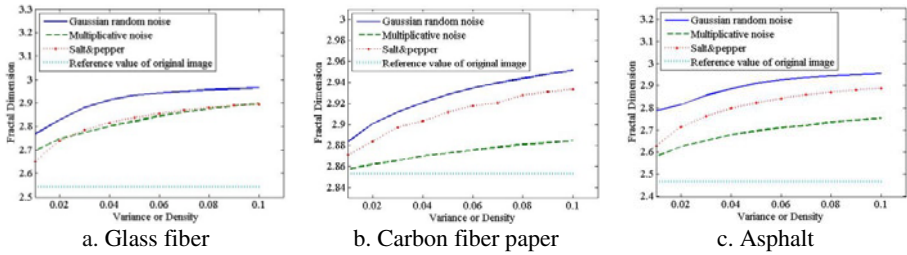


Fig. 1. The Effect of noise on fractal dimension

From Fig. 1, it can be concluded that all these three kinds of noises have a significant effect on fractal dimension, and therewith Gaussian random noise shows the most significant effect. And the fractal dimension increases with the increasing density or invariance of these three noises. As far as different images concerned, carbon fiber paper image suffers little from the noises. Taking salt and pepper noise for example, during the computation via DBC algorithm, when box with the size of more than 2 pixels is covered to an image, the extremum deviation value of gray levels of this region obviously becomes larger. The extremum deviation increases with increasing density of salt and pepper noise, resulting in an increase in fractal dimension. Fig.1 exhibits this variation trend as well. From Fig. 1, it can be seen that among the three kinds of noises random Gaussian noise has the largest influence on fractal dimension, due to its most corrosive action on image data, which makes image roughest. In one word, the fractal dimension of material image changes a lot with the increase of variance or density of three noises, and the performance of DBC algorithm is not satisfying in the case of noise pollution. Thus, it is essential to improve DBC algorithm.

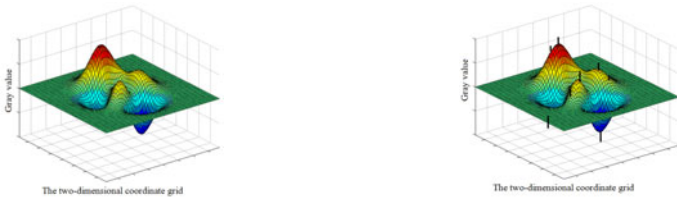
Fractal dimension calculation is based on gray-level of pixel, the accuracy of fractal dimension is very important for a quantitative description of the object. Generally, image preprocessing (such as smoothing) is used to denoise. Here, the conventional DBC algorithm is improved on the base of noise characteristics, and no image preprocessing is needed. In the follow sections, details on this anti-noise DBC algorithm will be discussed and comparison with conventional method will be made.

3 The Principle of Anti-noise DBC Algorithm

As shown in Fig. 2a, the original DBC algorithm considers the contribution between the maximum and minimum gray level l and k in the $(i, j)^{th}$ grid, and box number of the $(i, j)^{th}$ grid is $l-k+1$. It can be seen that the limitation of this algorithm is that only the maximum and minimum pixels of the grid are used while other pixels in the grid give no contribution. Nipon[8] proposed a modified DBC algorithm which replaced maximum and minimum with second maximum and second minimum pixel gray value. Wang Juan[9] used the average of gray level within grid to compute fractal dimension and the improved $n_r(i, j)$ is

$$n_{ij} = \begin{cases} 1 + \omega \cdot (\bar{g}_{ij} - g_{\min}) & (\bar{g}_{ij} - g_{\min}) > (g_{\max} - \bar{g}_{ij}) \\ 1 + \omega \cdot (g_{\max} - \bar{g}_{ij}) & \text{else} \end{cases} \quad (3)$$

where ω is weight and if $(\bar{g}_{ij} - g_{\min}) > (g_{\max} - \bar{g}_{ij})$, $\omega = N / N_{\min}$, else $\omega = N / N_{\max}$. N is the total number of pixel within grid, N_{\min} is the number of pixel with value between the minimum and the mean value, and N_{\max} is the number of pixel with value between the mean value and the maximum. This algorithm takes each pixel within the grid into account. Since noise is added to the captured image, gray values of some pixels will change, and the maximum and minimum will also change, as shown in Fig. 2b. In this situation, the change of box number in the grid will result in a deviation. In one word, the DBC algorithm and above modified algorithm are both lacks of immunity to noise because the extremum deviation is very sensitive to noise.



a. The ideal state of fractal dimension calculation b. The situation with noise

Fig. 2. Comparison between origin image and polluted image

As we all know, standard deviation reflects the data dispersion and have a thick skin to noise. Standard deviation can be used to replace extremum deviation for a better capability of noise proof. Firstly, the pixels within grid are arranged in increasing sequence $g(n)$ according to the gray value. Secondly, find sequence $g'(n)$, and n should satisfy $\max\{1, \text{round}(LS^2)\} < n < \text{round}(US^2)$, where $0 \leq L < U < 1$. Thirdly, compute the standard deviation of $g'(n)$ and denote the box number as σ . Therefore, the box number of the $(i,j)^{th}$ grid is $2 \cdot \sigma + 1$. In other words, $2 \cdot \sigma$ takes the place of extremum deviation in the origin DBC algorithm. In addition, it's essential to scan the whole image in order to count the number of box in different grid scales. Generally speaking, the grid scale is from 2 to $M / 2$. However, if the grid scale is too small, it cannot reflect the statistical characteristics of the standard deviation, and thus the grid scale is modified as $\sqrt[3]{M} < r < M / 2$ [10].

4 Comparison between Anti-noise and Conventional DBC

By using anti-noise DBC algorithm, fractal dimensions of glass fiber image with different noises are computed. For comparison, the conventional DBC algorithm with a median filter preprocessing is also tested, and the results under Gaussian random noise, multiplicative noise and salt and pepper noise are shown in Fig. 3(a), Fig. 3(b) and Fig. 3(c), respectively. In the experiment, U is 0.75 and L is 0.25.

The variation of fractal dimension with the variance of Gaussian random noise is plotted in Fig. 3(a). The real line denotes the result of the conventional DBC algorithm, which gives a great deviation from the reference value computed by original image. The results of both anti-noise DBC algorithm and conventional DBC algorithm with a median filter preprocessing approach the reference value computed by original image, and both algorithms show a good denoising performance. When the noise variance is more than 0.02, anti-noise DBC algorithm gives a better fractal dimension calculation which value is closer to the reference value compared to the conventional DBC algorithm with preprocessing.

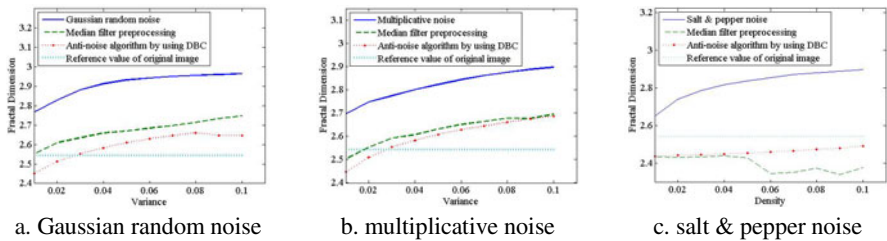


Fig. 3. The Effect of noise on fractal dimension

Fig. 3(b) shows the effect of multiplicative noise on the fractal dimension, which is similar to that of Fig. 4. Comparison to the conventional DBC algorithm with preprocessing, when the noise variance is more than 0.025, anti-noise DBC algorithm gives a better fractal dimension value. And the variance trend for anti-noise DBC algorithm is smooth.

Fig. 3(c) shows the variance of fractal dimension with the density of salt & pepper noise. Anti-noise DBC algorithm gives a better fractal dimension value in comparison with the conventional DBC algorithm with preprocessing.

From both Fig. 3(a) and Fig. 3(b), it can be seen that the calculation results of both anti-noise DBC algorithm and conventional DBC algorithm with a median filter preprocessing are great than the reference value in most noise variances, while in Fig. 3(c), the results of both algorithms are less than the reference value. It is due to the affect of anti-noise DBC algorithm and median filter preprocessing on different kinds of noise.

5 Conclusions

Fractal dimension gives a quantitative characterization on digital image, and its accuracy guarantees the veracity of texture analysis, image segmentation, shape description and characterization. While captured image is usually disturbed by noise, the fractal dimension is positively correlated with the noise variance or density. In 3D Euclid space, fractal dimension of image is a fraction between 2 and 3. Due to the disturbance of noise, the image became rougher, and its fractal dimension becomes larger, close to 3. It shows that fractal dimension description is consistent with the sensation of image roughness via the human visual system. Compared to salt & pepper

noise and multiplicative noise, random Gaussian noise has the most significant influence on fractal dimension under the same variance or density. Since the fractal dimension value calculated by conventional DBC algorithm gives a great deviation from the reference value computed by original image, an anti-noise differential box-counting algorithm based on noise characteristics is proposed, which takes full advantage of each pixel in the box and replaces the extremum deviation with standard deviation. Comparison between anti-noise DBC algorithm and conventional DBC algorithm with preprocessing shows that anti-noise DBC algorithm gives a better fractal dimension value in most case. And the corresponding variance trend of fractal dimension with variance or density is smooth. The proposed anti-noise DBC algorithm can be used to compute fractal dimension without preprocessing like smoothing and the computation results shows an effective denoising performance of this algorithm.

Acknowledgments. This paper is supported by the Project 51106116 supported by National Natural Science Foundation of China and the Fundamental Research Funds for the Central Universities.

References

1. Jayalalitha, G., Uthayakumar, R.: Recognition of Cervical Cancer Based on Fractal Dimension. In: International Conference on Advances in Recent Technologies in Communication and Computing, vol. 11, pp. 532–536 (2009)
2. Ling, L., Ming, L., YuMing, L.: Texture Classification and Segmentation Based on Bidimensional Empirical Mode Decomposition and Fractal Dimension. In: First International Workshop on Education Technology and Computer Science, vol. 5, pp. 574–577 (2009)
3. Dominkovics, C.: Dendrite material identification method using fractal analysis. International Spring Seminar on Information and Automation 5, 200–203 (2010)
4. Zhengyu, L., Yangyong, S.: Research on Dimension Algorithms about Fractal of Microstructure Curve of Rock by Using Box-Counting Dimension Method. In: International Conference on Intelligent System Design and Engineering Application (ISDEA), vol. 11, pp. 189–192 (2010)
5. Pentland, A.P.: Fractal-Based Description of Natural Scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 661–674 (1984)
6. Jun, K.: *Analysis on fractals*, pp. 25–33. Cambridge University Press, Cambridge (2001)
7. Nirupam, S., Chaudhuri, B.B.: An efficient differential box-counting approach to compute fractal dimension of image. *IEEE Transactions on Systems, Man, and Cybernetics Society* 24, 115–120 (1994)
8. Nipo, T.: Fractal dimension estimation using modified differential box-counting and its application to MSTAR target classification. In: *IEEE Intl Conf. on Systems, Man, and Cybernetics*, Hammamet, Tunisia (2002)
9. Juan, W., Lin-lin, C., Kang-ze, Y.: Efficient approach to compute fractal dimension in SAR images. *Computer Engineering and Design* 2, 315–325 (2005)
10. Zhi, Z., Fu-an, D., You-li, W.: Estimation of fractal dimension for 2-D gray image. *Computer Applications* 12, 2853–2867 (2005)
11. Wen-Li, L., Yung-Chang, C., Kai-Sheng, H.: Robust Calculation of Fractal Dimension of Images and its applications to classification of ultrasonic liver images and texture images. In: *IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 656–659 (2002)

A Spectral Matching for Shape Retrieval Using Pairwise Critical Points

Zhen Pan¹, Guoqiang Xiao^{1,*}, Kai Chen¹, and Zhenghao Li²

¹ College of Computer & Information Science, Southwest University,
Chongqing, 400715 China

² Laboratory of Optoelectronic Technology and Systems of the Education Ministry of China,
Chongqing University, Chongqing, 400044 China
{b222005080605275, gqxiao, chenkaimm}@swu.edu.cn,
lizhenghao@cqu.edu.cn

Abstract. The matching and retrieval of shapes is an important challenge in computer vision. A large number of shape similarity approaches have been developed. In this paper, we employ two approaches for improving shape retrieval. First, we use angle gradient to extract contour's critical points, this is a simple approach which decrease the computational cost while retain spatial information of shapes. Second, we present a pairwise similarity measure, which is a quadratic assignment problem. This problem is approximately solved by spectral technique. This method is tested on standard MPEG-7 shape database using the standard performance evaluation scheme. The experimental results indicate that the proposed method outperforms the closely relate method.

Keywords: shape retrieval, angle gradient, spectral technique.

1 Introduction

The ultimate aim of computer vision is to make an artificial system with the capability of human visual system. To this aim, object recognition is one of the fundamental problems in vision to be solved. An object has many different features like color, texture, shape etc. Shape is probably the most important feature of an object, which can be perceived. Therefore, the study of shape retrieval, which mainly utilizes the shape features of the objects to perform the retrieval task, continues to be a crucial issue in computer vision. Shape retrieval has been studied extensively in the past. A mass of methods have been proposed for shape similarity measure. The existing techniques can be generally classified into two categories: contour-based methods and region-based methods. The contour-based methods, such as the curvature scale space(CSS) [2], which is based on finding points of inflection on the curve at various levels of detail, and is invariant under the affine transform. Fourier-based descriptors [3], which are easy to be implemented are based on the well-developed theory of Fourier analysis. Shape context [4] is a local method for describing shapes. Each point is represented by a histogram of distributions of position of other points with respect to

* Corresponding author.

this reference point in the log-polar space. Attalla and Siy [5] proposed a polygonal shape descriptor at multi-resolution for shape matching. The measure of shape similarity is usually a metric distance between the acquired feature vectors.

On the other hand, in region-based methods, the global information of all the pixels within a shape is taken into account to obtain the shape representation, such as Moment shape descriptors, which are usually concise, robust, and easy to compute and match. These include geometric moments, Legendre moments, Zernike moments and pseudo Zernike moments [6]. However, the disadvantage of moment methods is that it is difficult to correlate high order moments with the shape's physical features.

In this paper, we used a contour-based method to shape retrieval. First, we propose a simple method to detect the critical point of the curvature, then we present a shape matching scheme. It provides a different framework, which is based on optimizing geometric pairwise affinities by solving a quadratic assignment problem. As the solution to the resulting computational problem is *np-hard*, it is approximated by spectral technique. Finally, we verify the proposed scheme by applying them to shape retrieval using the MPEG-7 database.

2 Angle Gradient

In this section, we describe how to extract the contour's critical points which are used for the match scheme. Analyzing the contour of an object is the first and most important step in shape matching. Essentially, the silhouette of object is interpreted as a continuous curve with perceived sense, while the perception is based on a certain scale under observation [7]. In this paper, the input to the system is a closed planar contour. But this contour is composed of too many points. Let C denote a shape contour, every contour is sampled with the same number of points use the method in Ref. [4]. The resultant points on the contour can be represented as $C = \{c_1, c_2, \dots, c_n\}, c_i \in \mathbb{R}^2$, where $c_i = (x_i, y_i)$ and n denotes the number of points on the contour.

Since different objects can be represented by different number of points, which are called contour's critical points. As a matter of fact, contour's critical points are usually the corner points of object, which are the visual part of shape. We usually consider that curvature extreme points close contact with visual perception. Nevertheless, as the result of the diversity of shape, simple shapes have few curvature extreme points with loss of important information. While complex shapes with lots of curvature extreme points increase computational cost. So just use the curvature extreme points will lead to big deviation. In order to decrease the computational cost while retain spatial information of the shape, we proposed a simple extreme points detection method to extract critical points. Suppose $\mu(i) = |gra_{i-1,i} - gra_{i,i+1}|$, where $gra_{i-1,i}$ represents angle between $c_{i-1}c_i$ and x-axis and $gra_{i,i+1}$ represents angle between $c_i c_{i+1}$ and x-axis. $\mu(i)$ is angle gradient of the i th point. If $\mu(i) < \lambda$, we think c_{i-1}, c_i, c_{i+1} are approximately in a straight line, which is to say $\mu(i)$ is small enough so that c_i can be eliminated. Otherwise c_i is retained. After eliminating these points, the result of the contour's critical points are summarized in Fig.1.

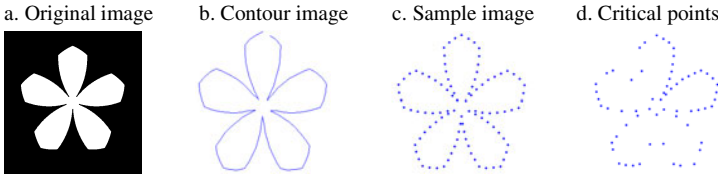


Fig. 1. Procedures and results of contour points extracted and sampled

3 Measure with Spectral Technique

In this section, we will use a different class of similarity measure computed over pairs of model points in an image. This will allow the similarity maximization to be expressed as a Quadratic Assignment Problem.

Given two shapes S_1 and S_2 and their sets of critical points $P = \{p_i\}_{i=1}^{n_1}$ and $Q = \{q_i\}_{i=1}^{n_2}$. Essentially, solving shape matching problem is equivalent to solving the matching of the corresponding points between P and Q . The correspondence is a mapping function σ indicating that p_i corresponds to $q_{\sigma(i)}$. To reduce notational clutter we will sometimes abbreviate $\sigma(i)$ as i' , so σ maps p_i to $q_{i'}$.

It is common to express the problem of finding a good correspondence as minimize a cost function, such as the follow cost function.

$$H(S_1, S_2; \sigma) = \sum_i C(i, i'). \tag{1}$$

Where $C(i, j)$ is the cost of matching $p_i \in P$ to $q_j \in Q$ based on some shape descriptors discrepancy measure. Consequently, the mapping function σ is the one that optimizes the cost. We use a different shape matching model computed over pairs of critical points on a shape. This will allow the similarity maximization to be expressed as a Quadratic Assignment Problem (QAP).

$$H(S_1, S_2; \sigma) = \sum_i \sum_j \Phi(d(i, j), d(i', j')). \tag{2}$$

Where $\Phi(d(i, j), d(k, l))$ is the similarity between critical points $p_i, p_j \in P$ and $q_k, q_l \in Q$. Optimizing (2) is denoted as the QAP, and its solution is known to be *np-hard*. We use spectral technique [9] to solve this problem. In order to solve (2), we follow a two step process that results in good solution to our problem.

A. Rewrite the Similarity Scheme (2)

For each pairs of assignments (a, b) , where $a = (i, i')$ $p_i, p_{i'} \in P$ and $b = (j, j')$ $q_j, q_{j'} \in Q$, there is an affinity that measure how compatible the points (p_i, p_j) are with the points (q_i, q_j) . We store the affinities on every pairs of assignment in the $n \times n$ affinity matrix M as follows:

1). $M(a, a)$ is the affinity at the level of individual assignments $a = (i, i')$. It measures how well the shape descriptor of p_i matches the shape descriptor of $q_{i'}$. Assignments that are unlikely to be correct (due to large distance between the descriptors) will be filtered out.

2). $M(a, b)$ describes how well the relative pairs of points $(q_{i'}, q_{j'})$ is preserved after putting them in correspondence with the pairwise points (p_i, p_j) . Here $a = (i, i')$ and $b = (j, j')$. If the two assignments do not agree (e.g. the deformation between (i, i') and (j, j') is too large) or if they are incompatible based on the mapping constraints (e.g. $i = j$ and $i' \neq j'$), we set $M(a, b) = 0$.

3). We require these affinities to be non-negative, symmetric and increasing with the quality of the match, without any loss of generality. M is an $n \times n$ positive matrix where $n = n_1 \times n_2$. As a and b are not numerals, and can not serve as matrix subscripts, we replace them with s and t , where $s = (i-1)n_2 + i'$ and $t = (j-1)n_1 + j'$.

$$M(a, b) = M(s, t) = M((i-1)n_2 + i', (j-1)n_1 + j') = \Phi(d(i, j), d(i', j')) \tag{3}$$

$$i, j = 1, \dots, n_1, i', j' = 1, \dots, n_2.$$

The similarity Φ is given by the kernel

$$\Phi(d(i, j), d(i', j')) = \exp\left\{-\frac{(d_{i,j} - d_{i',j'})^2}{2d_{i,j}d_{i',j'}}\right\}. \tag{4}$$

We propose to use the inner geodesic distance between the points $(p_i, p_j) \in P$, as the pairwise geometric measure $d(i, j)$, as it yields an affinity measure Φ that optimizes intrinsic geometric consistency, rather than shape descriptors consistency[4][10].

The correspondence problem reduces now to finding the assignments (i, i') that maximization the similarity such that mapping constraints are met. The matching $\sigma: S_1 \rightarrow S_2$ is the one that maximization the pairwise shape matching model H , given in (2). We represent the assignment function σ by an assignment matrix $X \in \{0,1\}^{n_1 \times n_2}$, where an element $x_{i,i'} = 1$ indicates that $p_i \in P$ is matched to $q_{i'} \in Q$. The matrix X is vectorized row-wise into $x \in \{0,1\}^{n_1 n_2}$. Using the above-mentioned notation, formula (2) can be rewritten as

$$H(S_1, S_2; \sigma) = x^T M x. \tag{5}$$

However, the matrix M is of dimension $n_1 n_2 \times n_1 n_2$, and matching the model would require solving huge linear programming problems or spectral decompositions. To resolve that, we propose to utilize shape context to reduce the number of candidate assignment of each points in S_1 . For each point $p_i \in P$, we keep the K points

$\{q_j\}_{j=1}^K \in Q$ whose shape contexts are the $K - NN$ features to $p_i \in P$. In this case the size of M and the dimension of the problem search space are considerably reduced.

B. Solve the QAP

The optimal solution x^* is the binary vector that maximization the similarity, given the mapping constraints:

$$x^* = \arg \max_x (x^T M x), \quad x \in \{0,1\}^{n_1 K}. \tag{6}$$

As this assignment problem is known to be *np-hard*, it is approximately solved. So relax the binary constraint and the matching constraints, such that the elements of x can take real values in $[0,1]$.

$$z^* = \arg \max_z z^T M z, \quad z^T z = 1. \tag{7}$$

Since M is symmetric, according to the the Rayleigh-Rit theorem [14]:

$$\lambda_{\max} = \max_{z^T z = 1} z^T M z. \tag{8}$$

And z^* is the eigenvector corresponds to λ_{\max} . By Perron-Frobenius theorem[15], since M is non-negative, z^* is strictly positive, finally, extract a binary approximation from the continuous solution by solving:

$$x^* = \arg \min_x z^{*T} x. \tag{9}$$

Equation (9) corresponds to a rectangular linear assignment problem. Therefore, given two shapes S_1 and S_2 their similarity is given by the optimal quadratic assignment score. Given the matching vector x the similarity measure is given by:

$$H(S_1, S_2) = x^{*T} M x^*. \tag{10}$$

The similarity between the two shapes S_1 and S_2 are measured by $H(S_1, S_2)$, which can be used for shape retrieval and shape matching.

4 Experimental Results

In this section we test our method for shape retrieval on the widely used shape databases, MPEG-7. We compare the results with six closely related methods for the shape retrieval. MPEG-7 shape database [8] contains 1400 shapes from 70 different categories (Fig.2) that they come from our side. It has been extensively used to evaluate the performance of many algorithms for shape matching and retrieval in the literature. This database is challenging because of interclass variations and also because of similarity between shapes in different classes.



Fig. 2. MPEG-7 shape database which contains 70 categories with 20 shapes each

The measure was taken to evaluate the performance of the different methods with respect to shape retrieval. One measure is precision and recall pair. Precision P is defined as $P=r/n$ and recall R is defined as $R=r/m$, where n is the total number of the retrieval shapes, r is the number of relevant shapes, i.e. truly similar to the query shape among the n retrieved shapes, and m is the total number of relevant shapes in the whole databases. Precision measures the accuracy of the retrieval and the speed of the recall and recall measures the robustness of the retrieval. It must be noted that since each shape in the database is taken as a query, the precision value of a specific recall is the average of the precision value of all the queries.

Another measure is retrieval accuracy. The normal test for the recognition accuracy is bulls-eye test [15]. Each shape in the database is taken as a query and matched to all the shapes in the database(including itself). The total number of similar shapes(which belong to the same class) of all the queries in the top 40 matches is counted as T . Since the maximum number of correct matches for each query is 20, the total maximum of correct matches is 1400×20 . The retrieval accuracy is then defined as $T/28000$.

In our experiment, the number of the initial points was set at 100. The angle gradient scale λ is set at 3° . The number of nearest neighbor features for each point used is 5. The retrieval accuracies for the six competing methods on MPEG-7 database are show in Table 1. From Table 1, we can see that among all the competing methods, our methods achieve the best score on the MPEG-7 database. It is noticeable that the retrieval accuracy of the proposed method is higher than other methods.

Table 1. Retrieval accuracies for the MPEG-7 databases

Method	Score (%)
<i>Shape Context</i> [4]	76.51
<i>Inner-distance</i> [10]	85.4
<i>Data-driven EM</i> [11]	80.03
<i>Contour flexibility</i> [12]	89.31
<i>Locally constrained diffusion</i> [13]	93.32
<i>CPDH</i> [1]	76.56
<i>Proposed method</i>	93.41

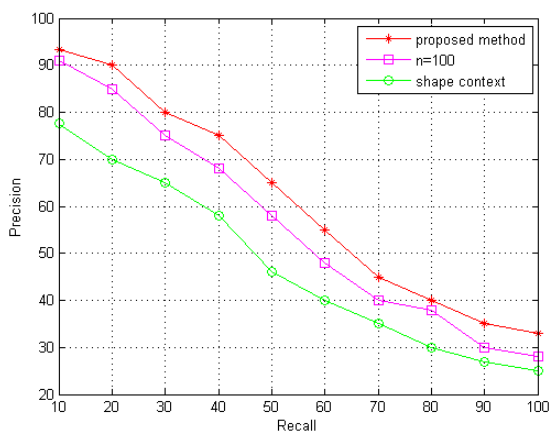


Fig. 3. Precision-Recall plot on the MPEG-7 database

Fig.3 illustrate the precision-recall plots, where each competing method is represented by a curve, for the MPEG-7 database. From precision-recall plots, we can see that shape retrieval with critical points has a higher precision than that with same points($n=100$). In addition to this, when recall value achieves 65%, the precision value of the proposed method on the MPEG-7 database is 18.77% higher than the traditional method shape context. All the experiment results indicate that the proposed method has significantly higher performance than that with same points or linear assignment method and is very suitable for shape retrieval.

5 Conclusion

This paper proposed a simple approach to detect critical points of shape and presents a pairwise framework for shape retrieval. First, we represent the shape with finite points set which was close contact with visual perception. Second, we proposed a different framework for shape similarity measure, and approximated the solution of a quadratic assignment problem using an efficient spectral technique. By applying it to the MPEG-7 database, we show that the proposed method outperforms closely relate methods.

References

1. Shu, X., Wu, X.: A Novel Contour Descriptor For 2D Shape Matching and its Application to Image Retrieval. *Image and vision Computing* 29, 286–294 (2011)
2. Abbasi, S., Mokhtarian, F., Kittler, J.: Curvature Scale Space Image in Shape Similarity Retrieval. *Multimedia Systems* 7, 467–476 (1999)
3. Kauppinen, H., Seppanen, T., Pietikainen, M.: An Experimental Comparison of Autoregressive and Fourier-based Descriptor in 2D Shape Classification. *IEEE Trans. Pattern Anal. Mach. Intell.* 17, 201–207 (1995)

4. Belongie, S., Malik, J., Puzicha, J.: Shape Matching and Object Recognition Using Shape Contexts. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 509–522 (2002)
5. Attalla, E., Siy, P.: Robust Shape Similarity Retrieval Based on Contour Segmentation Polygonal Multiresolution and Elastic Matching. *Pattern Recognition* 38, 2229–2241 (2005)
6. Zhang, D., Lu, G.: Review of Shape Representation and Description Techniques. *Pattern Recognition* 37, 1–19 (2004)
7. Marr, D.: *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman, San Francisco (1982)
8. Latecki, L.J., Lakamper, R., Eckhardt, T.: Shape Descriptors for Non-rigid Shapes with a Single Closed Contour. In: *Computer Vision and Pattern Recognition*, vol. 1, pp. 424–429 (2000)
9. Leordeanu, M., Hebert, M.: A Spectral Technique for Correspondence Problems Using Pairwise Constraints. In: *International Conf. Computer Vision*, vol. 2, pp. 1482–1489 (2005)
10. Ling, H., Jacobs, D.W.: Shape Classification Using the Inner-Distance. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 286–299 (2007)
11. Tu, Z., Zheng, S., Yuille, A.: Shape Matching and Registration by Data-driven EM. In: *Computer Vision and Image Understanding*, vol. 109, pp. 290–304 (2008)
12. Xu, C.J., Liu, Z., Tang, X.: 2D Shape Matching by Contour Flexibility. *IEEE Trans. on Pattern Anal. Mach. Intell.* 31, 180–186 (2009)
13. Yang, X., Koknar-tezel, S., Latecki, L.J.: Locally Constrained Diffusion Process on Locally Densified Distance Spaces with Application to Shape Retrieval. In: *Computer Vision and Pattern Recognition*, pp. 357–364 (2009)
14. Da, Z.: *Matrix Analysis and Applications*, pp. 528–530. Tsinghua University Press, Beijing (2004)
15. http://en.wikipedia.org/wiki/Perron%E2%80%93Frobenius_theorem#References

Skew Detection of Fabric Images Based on Edge Detection and Projection Profile Analysis

Zhoufeng Liu, Jie Huang, and Chunlei Li

School of Electric & Information Engineering, Zhongyuan University of Technology,
Zhengzhou 450007
{lzf, huangjie}@zzti.edu.cn, lichunlei1979@cse.buaa.edu.cn

Abstract. In this paper, a skew detection of fabric images scheme based on morphological method and projection profile analysis is proposed. The original image is convoluted using Sobel operator, and the weft boundary is generated by binarizing the convoluted image according to setting optimal threshold. A projection profile is computed at each angle, and the angle that maximizes a criterion function is regarded as the skew angle. Experimental results show that the skew angles detected by the proposed detection algorithm with higher accuracy.

Keywords: Fabric images, skew detection, sobel operator, projection profile analysis.

1 Introduction

In the processes of textile manufacturing, sample placed and other reasons will cause skew for the fabric, and the skew angle of fabric image has a negative influence on automatic analysis and recognition of fabric image. Therefore, it is necessary to develop a novel method for detecting skew angle, and can automatically correct fabric skew. Skew detection is often applied in document image processing. At present, many methods have been proposed, and they can be classified into three categories:

- The nearest-neighbor clustering method [1, 2] consists to calculate the line orientation that linked each two components related neighbor. And a histogram of the angles is formed to determine the skew angle. However, it is very time-consuming, moreover, it is sensitive to noise.
- Hough transform has been successfully applied to detect the lines and curves in the image. It can also be used to detect skew angle. In [3, 4], it has also been used for document image skew detection, Zhang [5] extended document skew detection to fabric skew detection. Its disadvantage is that one has to first extract text regions in the document image before the Hough transform is applied. This is not trivial for document images whose layouts are complex. Moreover, Hough transform-based methods are also computationally expensive.
- The third is projection profile (PJ) based approach for skew detection, and it was first proposed by Postl et al. [6]. The input document is rotated through a range of angles and a projection profile is calculated at each angle. Features are then extracted from each projection profile to determine the skew angle. This method

is very easy to be implemented, and it is suitable for simple documents. Recently, Li et al. [7] improved it using wavelet decompositions.

Different from document images, fabric images not only have rich texture, but also have fixed interval of parallel lines. In order to achieve better detection accuracy, taking advantage of the own characters of fabric images, a novel skew detection method is proposed based on edge detection and projection profile analysis in this paper.

2 The Proposed Fabric Skew Detection Method

As we know, fabric images are a special kind of texture images, whose warp and weft yarns are arranged closely and interlaced in specified organization, as shown in Fig.1 (a). Weft of these fabric images are arranged uniformly in fabric weaving process. Therefore, skew weft can represent the skew of fabric images. In this paper, we only detect the skew angle of fabric images by analysis of weft direction information.

The proposed method is based on edge detection and projection profile analysis. The edges of original fabric image are extracted using sobel operator. And the generated edge image is rotated through a range of angles with small steps and a projection profile is calculated at each angle. The skew angle is estimated by using our proposed criterion function. To speed up detection process and improve detection accuracy, our method is firstly performed in a coarse-to-fine mode, and then is performed in a detailed mode.

2.1 Edge Detection

For the fabric images, if we directly rotated each pixel through a range of angles, the computational complexity is very high; so, it is not suitable for on-line implementation. Moreover, the detection result is affected by gray value, noise, warp and etc. Therefore, pre-processing method is adopted to extract the useful information. In this paper, we only extract the weft boundary based on Sobel operator. Generally, the skew angle of fabric images is in $[-5^{\circ} \sim 5^{\circ}]$. Then, we use horizontal operator to extract weft boundary. And the operator is depicted as follows.

$$G_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad (1)$$

Firstly, the original image is convoluted using the above horizontal operator, and a convoluted image C is generated. Next, we set threshold to binarize the convoluted image, convert the binary image into one-dimension vector V , sort it by descend order, and note the ordered vector noted as SV . The binarization equation can be written as follows.

$$BW(i, j) = \begin{cases} 0, & C(i, j) > SV(f \times N) \\ 1, & otherwise \end{cases} \quad (2)$$

where N is the number of the total pixels, $f \in [0, 1]$ represents the threshold factor. According to selecting different factors, the first $f \times N$ pixels are set to 1, and they represent the edge boundary, as shown in Fig.1(b-d). From these figures, we can conclude that the boundary is clearer with the increase of factors f . Taking into account of detection accuracy and implementation efficiency, we select the factor f as 0.3.

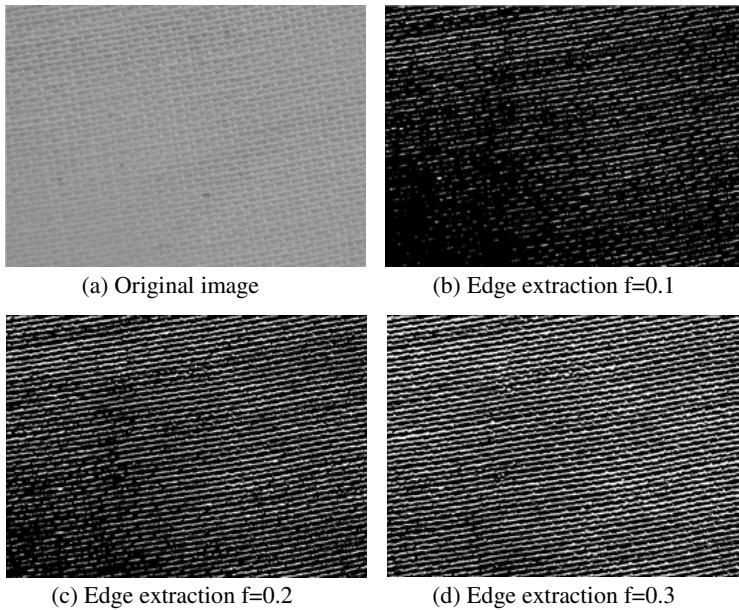


Fig. 1. Skewed plain cloth image

2.2 Projection Profile Analysis

A popular method for skew detection uses horizontal projection profile because the extracted edges are aligned along horizontal lines. The horizontal projection profile is applied on an image with size $M \times N$, then a column vector of size $L \times 1$ is obtained, where $L = \max - \min$. Elements of this column vector are the sum of pixel with values equals '1' in each row of the image. For a skew image, as shown in Fig.2 (a), we project it into Y axis, and the projection image is shown in Fig.2 (c). From this figure, we can see that the different value of projection count is small. When the skewed image is rotated with a given angle, let Y be perpendicular to the weft, as shown in Fig.2 (b), then we project the rotated edge image into Y axis, and the projection image is shown in Fig.2 (d). Because the Y axis is perpendicular to the weft, the difference value of projection count is larger. Based on this fact, during the detection process, we can rotate the edge image through $[-a, +a]$ with a small step,

and horizontal projection is computed at each angle, the rotated equation is written as follows.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{3}$$

As above described, only when Y is perpendicular to the weft, the difference value of projection count is largest. Therefore, we use the absolute of the variance of projection count as the criterion function. Because the center value can better represent their feature, we only select the center value which is in $[(\min + \max) / 2 - p, (\min + \max) / 2 + p]$, as shown in Fig.2 (e) and Fig.2 (f). In this paper, we select p as 100. Then, the criterion function can be written as follows.

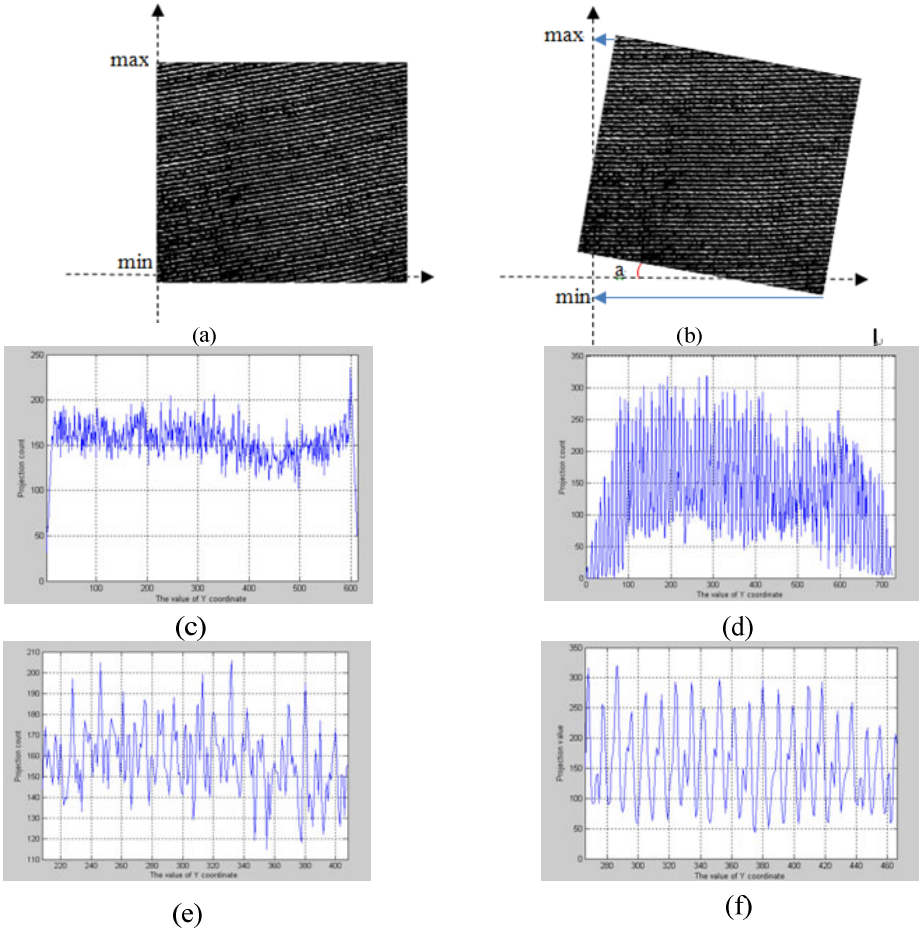


Fig. 2. Projection profile analysis

$$sum = \sum_{i \in [(\min+\max)/2-p, (\min+\max)/2+p]} abs(v_i - \bar{v}) \quad (4)$$

where v_i is the projection value, and \bar{v} is their mean value.

Because when Y is perpendicular to the weft, the difference value of projection count is largest. Therefore, the angle that maximizes this criterion function is regarded as the estimated skew angle. To speed up the search process, our method is performed in a coarse-to-fine mode. First, the search step size is 2^0 to get a coarse estimate. Noted the founding optimal angle as A . Then the search space is changed from $(A-2)^0$ to $(A+2)^0$ with a step size 0.1^0 .

3 Experimental Results

To evaluate our proposed method, we rotate the horizontal image by the given angle, and use the difference between the estimated and actual angles as evaluation criteria. In our experiment, we apply Photoshop software to generate a series of rotated image under different angle $[-10^0 \sim 10^0]$ with step size 2^0 . Table1 give the experimental results. It can be seen from the table that the average error of the algorithm is 0.15, and the average accuracy rate is 96%. For a given degree, the detection variance is no more than 0.5.

Table 1. The skew angle detection

Actual angle (0)	Detected angle (0)	Error degree(0)	Accuracy (%)
1	0.8	0.2	80%
3	3.3	0.3	90%
5	5.2	0.2	96%
7	7.1	0.1	98.6%
9	9.0	0	100%
-1	-0.9	0.1	90%
-3	-2.7	0.3	90%
-5	-5.4	0.4	92%
-7	-6.9	0.1	98.6%
-9	-9	0	100%

4 Conclusions

The skew angle of fabric image has a negative influence on automatic analysis and recognition of fabric image. In this paper, a skew detection of fabric images scheme based on morphological method and projection profile analysis is proposed. Firstly, Sobel operator is adopted to extract weft extraction, and then morphological method is used to reduce serrated points around weft boundary. In the end, a projection profile is computed at each angle, and the angle that maximizes a criterion function is

regarded as the skew angle. Experimental results show that the skew angles detected by the proposed detection algorithm with higher accuracy. Hierarchy detection scheme greatly reduce the execution time. Thus, the improved detection algorithm is capable of solving the skew problem in practical applications.

Acknowledgement. This work was supported by the National Natural Science Foundation of China, Grant No.61071175, and the key science and technology projects of Zhengzhou city, Grant No. 10PTGG382-1.

References

1. Lu, Y., Tan, C.L.: A nearest-neighbor chain based approach to skew estimation in document images. *Pattern Recognition Letter*, 2315–2323 (2003)
2. Liolios, N., Fakotakis, N., Kokkinakis, G.: Improved document skew detection based on text line connected component clustering. In: *Proc. Internat. Conf. on Image Processing*, Thessaloniki, Greece, pp. 23–27 (2001)
3. Amin, A., Fischer, S.: A document skew detection method using the Hough transforms. *Pattern Anal. Appl*, 243–253 (2000)
4. Khorissi, N., Namane, A., Mellit, A., et al.: Application of the wavelet and the Hough transform for detecting the skew angle in Arabic printed documents. In: *9th International Symposium on Signal Processing and Its Applications*, pp. 1–4 (2007)
5. Postl, W.: Detection of linear oblique structures and skew scan in digitized documents. In: *8th International Conference on Pattern Recognition*, France, pp. 687–689 (1986)
6. Zhang, R.L., Hu, Y., et al.: Skew Detection and Correction Method of Fabric Images Based on Hough Transform. In: *Second International Conference on Intelligent Computation Technology and Automation*, Changsha, Hunan, pp. 340–343 (2009)
7. Li, S.T., Shen, Q.H., Sun, J.: Skew detection using wavelet decomposition and projection profile analysis. *Pattern Recognition Letters*, 555–562 (2007)

Improved Algorithm of Separation and Identification of Touching Kernels and Foreign Materials in Digital Images

Zhining Liu and Lei Yan

Department of Automation, Beijing Forestry University, Beijing, China
Liu_lawrence@yahoo.cn, mark_yanlei@bjfu.edu.cn

Abstract. An improved method of separation and identification of touching kernels and foreign materials in digital images is proposed. The touching kernels are separated by using watershed algorithm based on morphological multiscale decomposition (MSD). Then, feature extraction of kernels is used for calculation of Mahalanobis distance. Finally foreign materials are identified by comparing Mahalanobis distance with the given threshold. The performance of the new algorithm is compared to that of a method based on a watershed algorithm and Mahalanobis distance (WMD). The experimental results showed that the efficiency of the proposed algorithm was superior with regard to WMD.

Keywords: Touching kernels, Foreign materials, Morphological multiscale, Watershed, Mahalanobis distance.

1 Introduction

Standards for quality grading of grains is the ensurement for high-quality and low-price of grains. One of the standards is the rate of foreign materials. To measure it, winnowing and screening are the most common methods. However, due to factors like different working conditions and operation by individuals, grading results often turn out to be poorly objective along with low efficiency and accuracy. As computer vision technology is becoming widely utilized in the field of grain inspection and identification. Efforts have been spent to identify the damaged grain kernels and dockages[1], to inspect and identify rice varieties [2], and to classify the grains [3-4]. Most of these studies have utilized well-defined images of grain kernels that are placed separately while in reality they often touch each other or even overlap. The clustered grain kernels make the feature extraction and identification difficult. Therefore, the separation algorithms are needed to separate touching kernels.

Vincent and Beucher proposed the traditional watershed segmentation theory based on the concept of topographic representation of image intensity, however, it requires a great amount of computation and its sensation to noise can leads to serious over-segmentation[5]. Shatadal developed a mathematical algorithm to disconnect touching kernels. But its limitation is that it failed when connected grain kernels formed relatively long isthmuses[6]. Yan proposed an algorithm based on contour segments

and ellipse fitting. It can separate the touching kernels successfully to some extent, however, when the grain is not ellipse-like or the contour segments are too short, the algorithm cannot achieve good separation result[7].

This paper presented an improved algorithm of separation and identification for touching sound kernels and foreign materials in digital images. First, two groups of images, images of touching sound grain kernels mixed with foreign materials and images of standard grain kernels, were filtered and converted into binary images. Then MSD method[8] was introduced and applied to separate touching kernels and foreign materials. The next step was to extract the morphological and color features of every individual kernel. Then, the Mahalanobis distances between the features of the separated sound kernels mixed with foreign materials and that of the standard kernels were calculated to be compared with a given threshold. Finally, the foreign materials were identified if their Mahalanobis distances were greater than the threshold.

2 Materials and Methods

2.1 Imaging System

The images were acquired by a color camera mounted on a stand stably. For uniform diffuse lighting, a circular fluorescent tube was placed below the surface level of the sample placement platform of the light chamber. The background was a black board.

The images acquired were divided into two groups: One is standard grain kernels images, in which the kernels are sound and do not touch each other; the other is touching sound kernels mixed with foreign materials images, in which the sound kernels and foreign materials were situated randomly and touching instances existed.

2.2 Grain Samples and Foreign Materials

Five kinds of grain kernels (common rice, rough rice, brown rice, common barley and glutinous barley) utilized as samples in our experiments were shown in Figure 1.

Broken kernels of same kind, unsound kernels of same kind, kernels of other kind with similar morphological features, and materials with great differences in both morphological and color features were classified as foreign materials.



Fig. 1. Image examples of individual grain kernel: (a) common rice (b) rough rice (c) brown rice (d) common barley (e) glutinous barley

3 The Proposed Algorithm

3.1 Image Pre-processing

The pre-processing steps for the two groups of images are the same. First, the color image was input and converted into the grayscale image. Then the median filter was used to remove the noise in the grayscale image. Finally, the image was converted into the binary one. The pixels with gray values higher than the threshold were given the value of one as objects; others were given the value of zero as background.

3.2 Image Segmentation

The segmentation method used in this paper is MSD method. In the first step, the image was decomposed into size-specific scale of distinct morphological markers. Then the converse synthesis of scales reconstructed the kernel size and shape while preventing already separated regions from merging. The detailed procedure is:

Step 1. The original color image was converted into a binary one;

Step 2. Markers were derived by an erosion process that started at the binary image and erodes regions to their markers ($I_{\text{marker-1}}$, $I_{\text{marker-2}}$, ..., $I_{\text{marker-s}}$).

Step 3. Starting from the binary marker image $I_{\text{marker-s}}$, a one-pixel gap between objects of $I_{\text{marker-s}}$ was generated by the watershed algorithm. Fig. 2a shows the skiz image of one scale;

Step 4. Objects of $I_{\text{marker-s}}$ were dilated under the binary mask image $I_{\text{marker-s-1}}$ while prohibiting merging of expanding marker regions until $I_{\text{marker-s}}$ reached a steady state. This mask operation was called the white-skiz operation[9]. Fig. 2b shows the white-skiz image of one scale;

Step 5. The operations in step 3 and 4 are repeated until $s = 0$. The resulting image is acquired. Fig. 2c is the image of the binary form of resulting image.

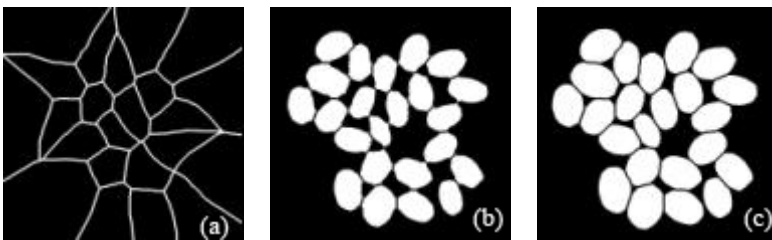


Fig. 2. Key images of segmentation process: (a) the skiz image of one scale (b)the white-skiz image of one scale; (c) the binary form of resulting image

3.3 Feature Extraction

Seventeen features are used in our study. Color features are the mean values of red, green and blue color bands of the kernels of the RGB colormap, the mean values of hue, saturation, value of the the kernels of the HSV colormap, and the standard deviation of

the hue. Morphological features include the area, length, width, major axis length, minor axis length, equivalent diameter, roundness(= (Perimeter)/(4 π ×Area)), aspect ratio(= major axis length/minor axis length), rectangular aspect ratio(= length/width), and extent(= area of the bounding box/area).

3.4 Mahalanobis Distance and Identification

Mahalanobis distance is a distance measure based on correlations between variables. It takes into account the correlations of the data set and is scale-invariant.

First the Mahalanobis distances based on the seventeen features between separated kernels and standard kernels are calculated. Then a threshold is determined through many observation tests. The threshold allows sufficient room for the feature variation in sound kernels and is far less than the Mahalanobis distances for any components suspected to be foreign materials. In the last step, each Mahalanobis distance is compared with the threshold, in which case the components with the value of Mahalanobis distance less than the threshold are identified as foreign materials.

4 Experimental Results and Discussion

4.1 Efficiency of Segmentation Algorithm for Different Kinds of Grains

The group of touching sound kernels and foreign materials included 409 of common rice, 467 of rough rice, 709 of brown rice, 401 of common barley, 213 of glutinous barley. Among them, the numbers of foreign materials are 20, 66, 86, 50, and 23, respectively. And the numbers of clusters in each kind are 65, 54, 75, 30, and 16, respectively. Fig.3 presents the color image example of rough rice, and the segmentation results processed by MW and MSD.



Fig. 3. Color image example and the segmentation results by MW and MSD; (a)color image example of rough rice; (b)segmentation result by MW; (c)segmentation result by MSD

MW's successful segmentation rates of common rice, rough rice, brown rice, common barley and glutinous barley are 83.96%, 82.06%, 82.42%, 89.17% and 94.37%. Its segmentation results are mostly affected by the touching degree and the internal markers. However, the MSD method turned out to be more efficient than MW, with the segmentation rates—99.29%、97.10%、97.78%、99.44% and 99.62%, respectively. MSD's disadvantages may lower the rates, like its strong dependence on

the size and shape of the structuring element, and its failure to separate small protrusions.

4.2 Efficiency of Identification Based on Mahalanobis Distance for Different Kinds of Grains

Images of five kinds of kernels mentioned above were processed by WMD and our algorithm. Fig. 4 shows the identification result of brown rice. The suspected foreign materials were labeled with crosshair signs.

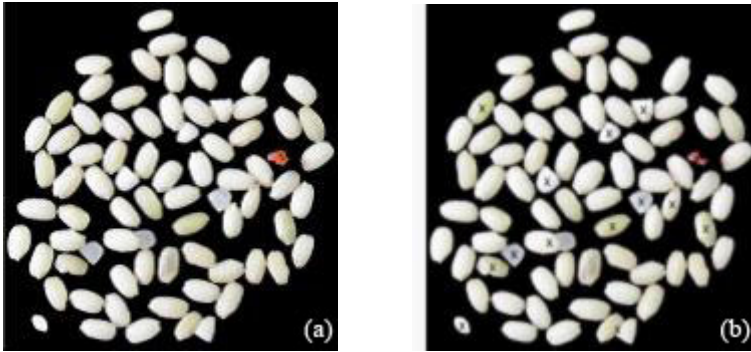


Fig. 4. Identification result of foreign materials: (a)Brown rice; (b)Identification result

By WMD, the identification accuracies were respectively 100.00%, 89.39%, 96.51%, 82.35%, and 100.00%, while by the proposed algorithm, they were respectively 90.00%, 96.97%, 100.00%, 90.20%, and 100.00%.

The results show that with regard to the efficiency of identification for foreign materials, the proposed algorithm works better than WMD. However, some factors give rise to the wrong identification for both methods, including the failure of segmentation and minor errors of the thresholds.

5 Conclusions

An improved algorithm of separation and identification for touching grain kernels and foreign materials is developed in this paper, in which MSD is first applied to the separation of grain kernels, and originally combined with Mahalanobis distance to identify the foreign materials in sound grain kernels. For performance evaluation, five kinds of grain kernels were used to conduct the experiments using WMD and our algorithm, respectively. The experimental results prove that the proposed algorithm is efficient and accurate in separating and identifying the touching sound kernels and foreign materials with a superior outcome over WMD.

Acknowledgments. We acknowledge the Fund for Young Scholars(2010BLX10), Beijing Forestry University, and the Fund of the 948 Project(2009-4-57), State Forestry Administration, P.R.China, for partial funding of the project.

References

1. Luo, X., Jayas, D.S., Symons, S.J.: Identification of damaged kernels in wheat using a colour machine vision system. *J. Cereal Sci.* 30, 49–59 (1999)
2. Majumdar, S., Jayas, D.S.: Classification of bulk samples of cereal grains using machine vision. *J. Agric. Eng. Res.* 73, 35–47 (1999)
3. Paliwal, J., Borhan, M.S., Jayas, D.S.: Classification of cereal grains using a flatbed scanner. *Can. Biosyst. Eng.* 46, 3.1–3.5 (2004a)
4. Carter, R.M., Yan, Y., Tomlins, K.: Digital imaging based classification and authentication of granular food products. *Meas. Sci. Technol.* 17(2), 235–240 (2006)
5. Vincent, L., Beucher, S.: The morphological approach to Segmentation: An introduction. School of Mines, Paris, France, Internai Rep. CMM (1989)
6. Shatadal, P., Jayas, D.S., Bulley, N.R.: Digital image analysis for software separation and classification of touching grains. I. Disconnect algorithm. *Trans. ASAE* 38, 635–643 (1995)
7. Yan, L., Park, C.-W.: New separation algorithm for touching grain kernels based on contour segments and ellipse fitting. *J. Zhejiang Univ.-Sci. C.* 12, 54–61 (2011)
8. Schmitt, O., Hasse, M.: Morphological multiscale decomposition of connected regions with emphasis on cell clusters. *J. Computer Vision and Image Understanding* 113, 188–201 (2009)
9. Preteux, E.: Watershed and Skeleton by Influence Zones: A Distance-Based Approach. *J. Mathematical imaging and Vision* 1, 239–255 (1992)

Object Detection Based on Multiclass Discriminative Field

Xiaofeng Zhang^{1,2}, Qiaoyu Sun¹, and Yue Lu¹

¹ Dept. of Computer Science and Technology,
East China Normal University, Shanghai, China

² School of Computer Science and Technology,
Nantong University, Nantong, China
ntuzxf@163.com,
qiaoyusun@163.com,
ylu@cs.ecnu.edu.cn

Abstract. In this paper, we present a novel object detection scheme that uses information of the sample fragments. These sample fragments are extracted by decomposition of the sample contour. Then, the candidate fragments corresponding to the sample fragments are detected from the images by partial Hausdorff distance. The Multiclass Discriminative Field (MDF) is used to select the most probable fragments from candidate fragments. The parameter estimation and inference of the MDF are simplified by using the candidate fragments as nodes of a graph. With these selected fragments, the contours of the objects can be obtained. The experiments on our postmark database and the ETHZ database show the feasibility of our proposed scheme.

Keywords: Object detection, Hausdorff distance, Multiclass Discriminative Field.

1 Introduction

Object detection is commonly used to separate interested objects from backgrounds. In the past years, it has raised a wide concern and interest because of its critical role in image retrieval, scene understanding and so on.

Features, such as contour, color and texture, are important for object detection. Fig. 1 shows some postmarks with different text and background. In this case, contour features have many advantages over others. For example, they are invariant to lighting conditions and object colors.

To detect an object, the sample contour is usually decomposed into several sample fragments [1]. Then, the candidate fragments of the images corresponding to the sample fragments are detected. Most methods employed Hough voting [2,3] to locate the centers of the contours by using the candidate fragments. And post processes have to be applied to select the fragments which belong to the contour of the object. In this work, we will directly select fragments by Multiclass Discriminative Field (MDF). MDF is one of Discriminative Random Field (DRF). It is usually used in multiclass labeling problems [4].

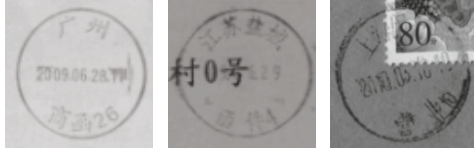


Fig. 1. Postmarks with different text and background

In our method, a sample fragment should be long enough and its Minimum Enclosing Rectangle (MER) should large enough. The sample fragments of an postmark are shown in Fig. 2. Then, we use a partial Hausdorff distance [5] method to detect the candidate fragments of the images corresponding to the sample fragments. Using the candidate fragments as nodes, a MDF is generated. The probability of each fragment belonging to an object will be inferred by Loopy Belief Propagation (LBP) [6]. Fragments with high probability are selected from candidate fragments and compose the contour of the object.

2 Candidate Fragment Detection

In order to detect the candidate fragments in the images, a partial Hausdorff distance is used.

Suppose $A = \{a_1, a_2, \dots, a_n\}$ is the point set of a sample fragment, $B = \{b_1, b_2, \dots, b_m\}$ is the point set of the edge of the image, and $B_i(x, y)$ centered point (x, y) is a part of B . B_i and A have a same MER. To detect whether B_i contains A , we use a partial Hausdorff distance [5]

$$h(A, B_i) = \frac{1}{N_a} \sum_{a \in A} d(a, B_i), \tag{1}$$

where $d(a, B_i) = \min_{b \in B_i} \|a - b\|$, $\|\cdot\|$ is some underlying norm (In this paper, it is Euclidean distance). If $h(A, B_i)$ is less than a threshold h_t , there will be a candidate fragment whose center is point (x, y) . Fig. 3(a) shows the candidate fragments detected by this partial Hausdorff distance method.

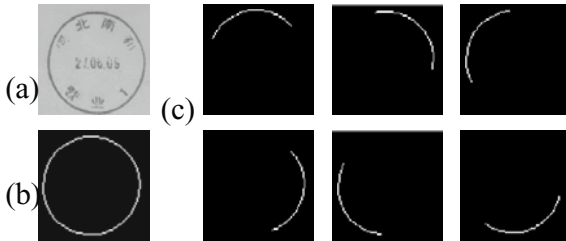


Fig. 2. Sample fragments of a postmark. (a) Sample image. (b) Sample contour. (c) Sample fragments.

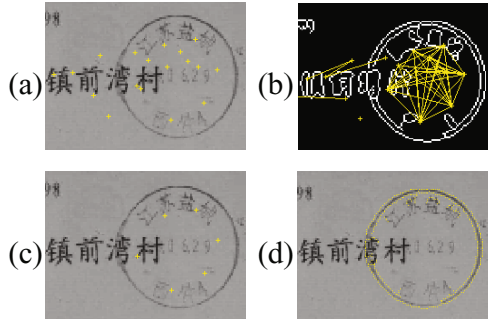


Fig. 3. An postmark detection process. (a) Candidate fragments detected by a partial Hausdorff distance. The crosses are centers of fragments. (b) An undirected graph for MDF. The crosses are the nodes of the graph and the straight lines are edges of the graph. (c) Fragments selected by MDF. (d) An postmark contour obtained by connecting the fragments.

3 Fragments Selection by MDF

Multiclass Discriminative Field is introduced for parts-based object detection [4]. In MDF, the distribution over all the labels x given the observations y can be written as

$$P(x|y) = \frac{1}{Z} \exp \left(\sum_{i \in S} A(x_i, y) + \sum_{i \in S} \sum_{j \in N_i} I(x_i, x_j, y) \right), \quad (2)$$

where Z is a normalizing constant known as the partition function, S is the set of all nodes, N_i is the neighbourhood of the node i , A and I , which called as association potential and interaction potential, are the unary and pairwise potentials, respectively.

The association potential A is

$$A(x_i, y) = \sum_{k=1}^C \delta(x_i = k) \log P'(x_i = k|y), \quad (3)$$

where C is the number of classes, $\delta(x_i = k)$ is 1 if $x_i = k$ and 0 otherwise, $P'(x_i = k|y)$ is the probability of x_i belonging to class k under the observation y .

The interaction potential I is

$$I(x_i, x_j, y) = \sum_{k=1}^C \sum_{l=1}^C \nu_{kl}^T \mu_{ij}(y) \delta(x_i = k) \delta(x_j = l), \quad (4)$$

where ν_{kl} are the model parameters and $\mu_{ij}(y)$ encodes the pairwise features.

Kumar et al. [4] constructed an undirected graph for MDF. The nodes of the undirected graph are candidate patches of the object. At the beginning of their method, the category which the node belongs to is uncertain, so MDF has a complicated parameter estimation period.

Using our candidate fragments as nodes, an undirected graph for MDF is generated (Fig. 3(b)). A node x_i contains the information of its location and the partial Hausdorff distance to the corresponding sample fragment. Each sample fragment decomposed from the sample contour is a category. The node belongs to the category which has the minimal partial Hausdorff distance. Because every node is known to a certain category, the parameter estimation is simplified.

The association potential A can be seen as a local decision term which decides the association of a given node to a certain category ignoring its neighbors. We describe P' using delta function [7]

$$P'(x_i = k|y) = \begin{cases} \frac{1}{2}(1 + \cos \frac{\pi H}{T_1}) & \text{if } H \leq T_1 \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

where H is a Hausdorff distance, T_1 is a threshold. The function approximately indicates the principle that smaller the Hausdorff distance is, higher probability the fragment belongs to the contour of an object.

The interaction potential I is seen as a data dependent smoothing function. $\mu_{ij}(y)$ is a pairwise feature, which indicates the similarity of the two neighbor nodes. According to [4], $\mu_{ij}(y)$ is the distance between centers of two candidate fragments i and j . ν_{kl} , which is a scale to measure $\mu_{ij}(y)$, is denoted as a function of the distance between sample fragment k and l . By the delta function, $\nu_{kl}^T \mu_{ij}(y)$ will be set as

$$\nu_{kl}^T \mu_{ij}(y) = \begin{cases} \frac{1}{2}(1 + \cos(\pi|D - 1|)) & \text{if } |D - 1| \leq T_2 \\ 0 & \text{otherwise} \end{cases}, \quad (6)$$

where $D = |D_{ij}/D_{kl}|$ denotes the difference between D_{ij} and D_{kl} , D_{ij} and D_{kl} are distance vectors, T_2 is a threshold. Similar to the association potential, smaller the difference between $\mu_{ij}(y)$ and ν_{kl} is, higher probability two nodes i and j belong to the same object.

With MDF model, we can calculate the marginal posterior probability of node x_i with label c_i . Suppose there are n sample fragments. $L = \{l_1, l_2, \dots, l_n\}$ is the set of label. Each node x_i corresponds to a sample fragment r if its label c_i equals to l_r . LBP is used to infer the marginal estimates and it is a message passing algorithm proposed by Pearl [6]. The nodes with large posterior probability are selected as a correct fragment of the object. Fig. 3(c) shows the fragments selected by MDF. The contour of the object in Fig. 3(d) composes of the fragments matching with the sample fragments.

4 Experiments

To evaluate the valid of the proposed method, we apply it to detect sample objects on our postmark database and the ETHZ database.

The postmark database has 120 envelope images which are categorised into three groups according to the complexities of their backgrounds. Each group has 40 images. The postmarks in the group (I) and (II) have similar sizes and clear contours, while the postmarks in the group (III), comparatively, have various sizes and false contours. Fig. 4 illustrates examples belonging to group (I), (II) and (III) respectively. We measure the performance of our method by *recall* and *precision*

$$\text{Recall} = \frac{\text{The number of correct detected objects}}{\text{The number of true objects}}, \quad (7)$$

$$\text{Precision} = \frac{\text{The number of correct detected objects}}{\text{The number of detected objects}}. \quad (8)$$

All postmarks of the first two groups are detected accurately with the *Recall* = 100% and *Precision* = 100%. In the last group, despite their complex backgrounds, most of the postmarks are detected. Its *Recall* = 95% and *Precision* = 55.6%. Some results with complex backgrounds are shown in Fig. 4. Fig. 4(a) illustrates some correct examples which locate the postmarks with proper rectangles. However, some other shapes like a part of a circle are mistaken for the contour of a postmark in Fig. 4(b).



Fig. 4. Some results of postmark detection. (a) Correct results. (b) Failed examples.

We also applied our method on ETHZ database [8]. Our algorithm is impotent to implement large rotations and scale changes. Hence we selected some proper images of this database to illustrate the results. Fig. 5 shows some results of apple logo and giraffe. Each row has a same scaling rate. Despite their deformation, backgrounds changing, color and texture variation, the rectangles cover the most parts of the objects.

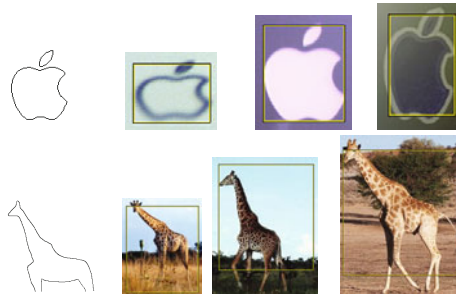


Fig. 5. Some results of ETHZ database. The first column is the sample contours provided by ETHZ database and the rest are the results of the images.

5 Conclusion

In this paper, an object detection method was proposed. We decompose a sample contour into sample fragments. With the application of the partial Hausdorff distance, candidate fragments corresponding to sample fragments are detected in the images. We simplify the parameter estimation and inference of MDF by using candidate fragment detection. With this MDF, proper fragments are selected. Finally, contours are obtained by connecting these candidate fragments. Based on contour decomposition and MDF, our model, which requires only a sample contour, tolerates some distortion, part of contour missing and various backgrounds. The experiments show that our method is a practicable scheme.

Acknowledgments. This work is partially supported by the Science and Technology Commission of Shanghai Municipality under research grant 09510701800.

References

1. Ferrari, V., Fevrier, L., Jurie, F., et al.: Groups of adjacent contour segments for object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 36–51 (2008)
2. Shotton, J., Blake, A., Cipolla, R.: Contour-based learning for object detection. In: *International Conference on Computer Vision*, pp. 503–510 (2005)
3. Ferrari, V., Jurie, F., Schmid, C.: From images to shape models for object detection. *International Journal of Computer Vision* 87, 284–303 (2010)
4. Kumar, S., Hebert, M.: Multiclass discriminative fields for parts-based object detection. In: *Snowbird Learning Workshop*, vol. 164 (2004)
5. Takacs, B.: Comparing face images using the modified hausdorff distance. *Pattern Recognition* 31, 1873–1881 (1998)
6. Pearl, J.: *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann (1988)
7. Zhao, H.K., Chan, T., Merriman, B., et al.: A variational level set approach to multiphase motion. *Journal of Computational Physics* 127, 179–195 (1996)
8. Ferrari, V., Tuytelaars, T., Van Gool, L.: Object Detection by Contour Segment Networks. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006, Part III*. LNCS, vol. 3953, pp. 14–28. Springer, Heidelberg (2006)

Co-processing Method for Automotive Vibration Signals on JBeam

Guofeng Qin¹, Minhu Fan¹, and Qiyang Li²

¹The CAD Research Center, Tongji University, Shanghai 200092, P.R. China
gfqing@yahoo.com.cn, fantj@foxmail.com

²The Chinese National CAD Application Technology Training Center for Engineering Design,
Shanghai 200092, P.R. China
qylcad@sina.com

Abstract. The characteristics of the Daubechies wavelet are studied. How to realize the Mallat algorithm and how to calculate the denoising threshold are discussed. A group of new application components on wavelet denoising are developed on JBeam platform. Finally, the components are utilized to denoise the real signal of bench test of automotive with sample cases, and the denoising effect of each Daubechies mother wavelet is analyzed to prove that the most suitable mother wavelet is Db14.

Keywords: JBeam, CEA, Daubechies wavelet, Mallat Algorithm, Vibration signal, Co-processing.

1 Introduction

In recent years, many researches have been done on wavelet transformation for signal denoising [1,2]. But how to choose mother wavelet and threshold of denoising achieving accurate denoise are not completely resolved [3]. Different kind of signal can fit less kind of mother wavelet well, so the choice of mother wavelet always changed for other signals. The choice of threshold is in the same situation, there is no method to find a threshold that can fit all situations. It is still a difficult problem that how to quantify the denoised wavelet coefficients [1, 4]. Donoho have given a hard threshold method and a soft threshold method [5]. Most studies only stay in the theoretical research. In this work, we will use the theory into practice and give an example to show how to choose the threshold and quantizing method with dbn wavelet.

2 Daubechies Wavelet and Threshold Denoising

2.1 Daubechies Wavelet

Daubechies wavelet belongs to discrete wavelet transform, is compactly supported orthonormal wavelet. In addition to db1, other Daubechies wavelets have no other analytical expressions. However, in practice, the following methods can be used to get the scale coefficients and wavelet coefficients [6].

$$P_M(z) = \sum_{k=0}^{N-1} \binom{N-1+k}{K} z^{N-1+k} \left(-\frac{1}{4}z^2 + \frac{1}{2}z - \frac{1}{4} \right)^k \tag{1}$$

Resolve the null point of $P_M(z)$, we take two null points of four complex null points of $P_M(z)$ and take one null point of two real null points of $P_M(z)$. The above formula will be rewritten to:

$$P_N(z) = [a_N \prod_{k=1}^K |v_k|^{-1} \prod_{j=1}^J z_j^{-2}]^{\frac{1}{2}} \times (\prod_{k=1}^K (z - v_k) \prod_{j=1}^J (z^2 - 2z \operatorname{Re} z_j + |z_j|^2)) \tag{2}$$

$$H(\omega) = \left[\frac{1}{2}(1 + e^{-j\omega}) \right]^N P_N(e^{-j\omega}) = \frac{1}{2} \sum_0^{2N-1} h_N(n) e^{-jn\omega} \tag{3}$$

Then $h_N(n)$ is Scale coefficient of Daubechies, we can use the scale coefficient to resolve coefficient of filter, and then we can execute wavelet transform.

In my component, I selected db4, db8, db10, db14 to verify the denoising effect.

2.2 Threshold Denoising

Threshold. Assuming that the noise level at all scales are same, we use robust method to estimate the level of noise [6, 7], and the optimized threshold is as follow.

$$\lambda = \frac{C_{mid}(t)}{0.6745} \sqrt{2 \operatorname{Log} N} \tag{4}$$

Threshold Denoising. Set the wavelet coefficients which have small absolute value to zero, and then reconstruct the signal to reach the goal of denoising. The wavelet coefficients commonly used denoising methods are two ways to quantify [5]: the first is the hard threshold method, denoised wavelet coefficients is shown as following.

$$\hat{W}_{j,k} = \begin{cases} W_{j,k} & |W_{j,k}| \geq \lambda \\ 0 & |W_{j,k}| < \lambda \end{cases} \tag{5}$$

The second is the soft threshold method, denoised wavelet coefficients is shown.

$$\hat{W}_{j,k} = \begin{cases} \operatorname{sgn}(W_{j,k}) \times (|W_{j,k}| - \lambda) & |W_{j,k}| \geq \lambda \\ 0 & |W_{j,k}| < \lambda \end{cases} \tag{6}$$

As hard threshold method is not continuous in λ , which make wavelet reconstruction signal oscillation easily, it is generally use soft threshold to quantify factors [7].

3 JBeam Structure and Working Principle Component Analysis

3.1 JBeam Introduction and Basic Data Types and Structure of the Internal

JBeam is data-processing software of German AMS Company. JBeam has internal interface of the software according with standards of ASAM CEA. JBeam provides six

external services for the preparation of that the components can communicate and collaborate with each other.

3.2 Code and Algorithm Complexity

The decomposition algorithm can be described as follow:

1. Get the original signal data, and the length of data.
2. Begin with traverse the array of signal data.
3. If the data to visit is in the edge of array, cycle extension method is used to extend the data to make sure the convoluting can work.
4. Convolute the signal data with the filter.
5. Write the result which is got by H filter in the front of the temporary array; and write the result that is got by G filter in the tail of the temporary array.
6. When the traversing finish, map the temporary array to original array.

Analyzing of the above algorithm, the program is completed in one cycle. The time complexity is $O(n)$. The space complexity is $O(n)$.

The denoise algorithm can be described as follow:

1. Get the decomposition data of signal.
2. Sort the first level of high frequency data.
3. Get the middle value and calculate the threshold by this value.
4. Begin with traverse the data of all high frequency.
5. If the data is smaller than the threshold, set the data as zero, otherwise, set the data as difference between data and threshold.
6. When the traversing finish, send the data array back.

The flow chart of denoise algorithm is as shown in Figure 1:

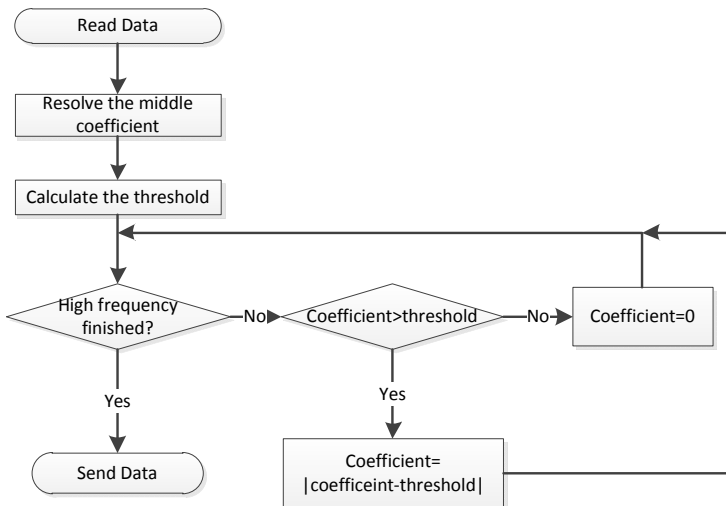
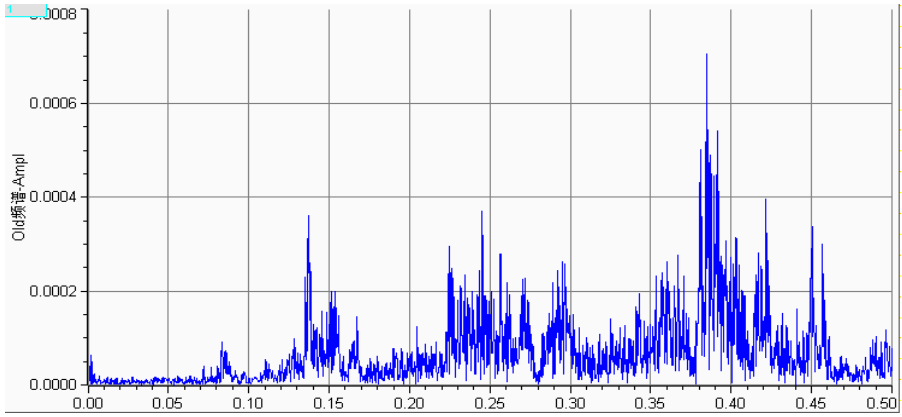


Fig. 1. Denoise algorithm

Analyzing of the chart, the most time spend on resolving the middle coefficient. So the time complexity is $O(n^2)$. The space complexity is $O(1)$.

4 Realistic Data Verification

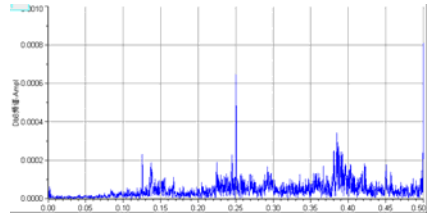
In a real test bed, a group of vibration signal sensors are used to collect the vibration signal data on the components of wavelet denoising. The results of signal process with a group of new developed application components can be seen as Figure 2.



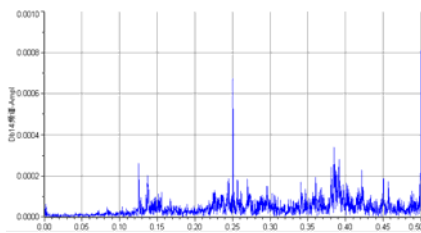
(a) Frequency spectrum of the original signal



(b) Frequency spectrum after Db4 denoising



(c) Frequency spectrum after Db8 denoising



(d) Frequency spectrum after Db10 denoising



(e) Frequency spectrum after Db14 denoising

Fig. 2. Frequency spectrum of the signal process

It can be seen from the above spectrum, the composition of noise signal have been reduced. The new signal can be used to extract the specific more easily, so Db14 is the best for this data.

5 Conclusion

The paper analyzes how to calculate threshold of the signal, and the algorithm to de-noise using the threshold. Then use a component to denoise the actual signal proved that the method in the paper is useful. But there still have some problems to solve. First, this paper can't calculate the threshold for signal with different noise level. Second, this component only used Daubechies wavelet function is not fit for all signals. The future workings should add other mother wavelet functions and continue to study the other threshold calculation in different cases.

Acknowledgement. We are appreciative of AMS GmbH to afford JBEAM software platform, which is helpful for our experiment and development.

References

- [1] Wang, B.-R., Yang, Y.-X., Cai, W., Yang, S.-Q.: Application of Wavelet Threshold De-noising Technique in Vibration Signals Processing. *Noise and Vibration Control* 28(6), 9–12 (2008)
- [2] Gao, Y.-H., Wang, W.-D., Jiang, Y.-S.: Selection of wavelet parameters in auto vibration signal denoising. *Journal of Hefei University of Technology (Natural Science)* 30(8), 975–979 (2007)
- [3] Mohideen, S.K., Perumal, S.A., Krishnan, N., Sathik, M.M., Kumar, T.C.R.: Image denoising multi-wavelet and threshold. *Computing, Communication and Networking* 24(2), 1–5 (2009)
- [4] Zhang, J.-X., Zhong, Q.-H., Dai, Y.-P.: The Determination of The Threshold and The Decomposition order in Threshold De-noising Method Based on Wavelet Transform. *Proceedings of The Chinese Society For Electrical Engineering* 24(2), 118–122 (2004)
- [5] Divid, L., Donoho: De-Noising by Soft-Thresholding. *IEEE Transaction on information Theory* 41(3), 613–627 (1995)
- [6] Yang, J.G.: *Wavelet Analysis and Its Engineering Applications*. China Machine Press (2005)
- [7] Wang, Z.-M., Dai, S.-J., Yue, H., Cai, H.-G.: Signal Denoising Based on the Wavelet Transform. *Journal of Hebei University of Technology* 29(5), 48–51 (2000)
- [8] Jiang, Y.-S., Wang, Q.-D.: De-noising Anlysis of Automobile Vibration Signal Based on Wavelet Transform. *Automobile Science & Technology* (4), 23–25 (2006)

A Method of Image Segmentation Based on Improved Adaptive Genetic Algorithm

Wenjiao Yu, Mengxing Huang, Donghai Zhu, and Xuegang Li

College of Information Science & Technology

Hainan University

Haikou, P.R. China

tsy19860303@yahoo.com.cn

Abstract. In order to quickly get the optimal threshold for image segmentation, a novel method of image segmentation based on improved adaptive genetic algorithm is presented, which adjusts the parameters of genetic algorithm adaptively according to the difference among the individuals and the diversity of population to ensure the convergence of the algorithm and avoid the precocious. The method combines with the Otsu method to get the best image segmentation threshold. The theoretically analysis and simulation experiments show that the threshold is more accurate and less time is consumed greatly by using the proposed method than Otsu image segmentation and other adaptive genetic algorithm.

Keywords: adaptation, genetic algorithm, Otsu method, optimal threshold, image segmentation.

1 Introduction

Image segmentation is one of the basic problems of the image processing and machine vision, its key point is: the image is divided into a number of sets that do not mutual overlapping zones; these zones either have meaning to currently mission or help to explain correspondence between them and the actual object or some parts of object [1].

In recent years, people raised some of partition techniques combined some specific theories, methods and tools with a number of new theories and methods of different disciplines has been proposed such as the use of Markov random field, mathematical morphology. Fuzzy ART, simulated annealing, genetic algorithm as well as the new segmentation algorithm is still emerging [2]. Among them, genetic algorithm has outstanding advantages of simple, robustness and parallelism, so it has gained great success in the application field [3]-[8].

Genetic algorithm is an iterative optimization algorithm to find global optimal segmentation threshold in the search space, so it is used to help determine the split threshold when make image segmentation. However, there are problems of poor astringency and premature in the basic Genetic Algorithm which cause great difficulties for image segmentation to find global optimal segmentation threshold. In order to improve convergence speed and resolve the problems of premature, Srinivas first proposed the concept of adaptive genetic algorithm in 1994 [11], from now on

adaptive crossover rate and mutation rate is applied to the genetic algorithm. Adaptive genetic algorithm, however, there is stagnation in the early evolution.[12] and [13] respectively to make improvements on this algorithm, but both of them only considering the differences of individual in population without considering the diversity of population, so there are further optimization. The differences among individuals in population and the diversity of population will be consider in this paper to adapter the genetic algorithm, so that to ensure convergence of the algorithm and to avoid premature, meanwhile combined with Otsu algorithm to calculate optimal threshold of image segmentation [9] [10].

2 Improved Adaptive Genetic Algorithm

2.1 Basic Adaptive Genetic Algorithm

In the genetic algorithm, the convergence speed of algorithm is directly affected by crossover rate and mutation rate. The size of the new individual in population is determined by crossover rate. The genetic algorithm with high crossover rate may destroy the excellent individual, but the genetic algorithm with low crossover rate may lead to premature. Mutation rate is the key factor of algorithm to jump out the local optimal solution. The genetic algorithm with low mutation rate couldn't generate new individuals, but on the other side the genetic algorithm will become a kind of pure random search algorithm with high mutation rate. In order to solve this problem, Srinivas and others proposed adaptive genetic algorithm. In order to maintain the diversity of population and ensure the convergence of genetic algorithm, the crossover rate and mutation rate can be automatically adjusted with the fitness. In the adaptive genetic algorithm, P_c and P_m are adaptively adjusted by the following formula:

$$P_c = \begin{cases} \frac{k_1(f_{\max} - f')}{f_{\max} - f_{avg}}, & f' \geq f_{avg} \\ k_2, & f' < f_{avg} \end{cases} \tag{1}$$

$$P_m = \begin{cases} \frac{k_3(f_{\max} - f)}{f_{\max} - f_{avg}}, & f \geq f_{avg} \\ k_4, & f < f_{avg} \end{cases} \tag{2}$$

where f_{\max} is the largest fitness in the population, f_{avg} is the average fitness of each generation in the population, f' is the larger fitness of the two individuals, f is the fitness of mutation individual, k_1, k_2, k_3, k_4 in the range (0,1).

We can find that the crossover rate and mutation rate will be smaller when the fitness of individual is closer to the maximum value by formula (1) and (2). Finally crossover probability and mutation probability is zero when their fitness are equal to the maximum value. However, this adjustment method is only suitable for the early

evolution of population because the best individuals will remain and do not change in the early of population, which likely lead to take local optimal solution as the global optimum. [12] and [13] improve this adaptive algorithm in different way to make sure that the crossover rate and mutation rate of individuals which have maximum fitness is not zero.

2.2 The Measurement of Population Diversity

The reduction of population diversity is another reason to the premature of genetic algorithm. Therefore, we need certain standards to measure the diversity of the population. When species diversity is high, the cross rate and the mutation rate of the population is set a smaller value. When the diversity of the population is low, the cross rate and the mutation rate of the population is set a bigger value. We can calculate the Diversity of Population (DOP) by formula (3).

$$DOP = \frac{1}{N} \sum_{i=1}^N \frac{|f_i - \bar{f}|}{f_{\max} - f_{\min}} \tag{3}$$

Where DOP is diversity of population, N is the size of the population, also is the number of individuals in population, f_i is the fitness of individual i , \bar{f} is the average fitness of the population, f_{\max} is the maximum fitness of the population, f_{\min} is the minimum fitness of the population. By formula (3), we can found that the DOP value is smaller when the populations are more concentrated, DOP values is bigger when population are more dispersed.

2.3 The Improved Adaptive Genetic Algorithm

By considering the difference between the individuals and the diversity of population to adjust the crossover rate and mutation rate of the population, so that to optimize the adaptive genetic algorithm. The improved crossover rate and mutation rate is calculated such as formula (4) and (5) below:

$$P_c = \begin{cases} P_{c1} * (1 - k_1 * DOP) * e^{-\frac{f' - f_{avg}}{f_{\max} - f_{avg}}}, & f' \geq f_{avg} \\ P_{c1} * (1 - k_1 * DOP), & f' < f_{avg} \end{cases} \tag{4}$$

$$P_m = \begin{cases} P_{m1} * (1 - k_2 * DOP) * e^{-\frac{(f_{\max} - f)}{f_{\max} - f_{avg}}}, & f \geq f_{avg} \\ P_{m1} * (1 - k_2 * DOP), & f < f_{avg} \end{cases} \tag{5}$$

Where f_{\max} is the largest fitness in the population, f_{avg} is the average fitness of each generation in the population, f' is the larger fitness of the two individuals,

f is the fitness of mutation individual, P_{c1} is the largest crossover rate, P_{m1} is the largest mutation rate, DOP is diversity of population, k_1, k_2 in the range (0,1).

3 The Combination of Improved Adaptive Genetic Algorithm and Otsu Method

Otsu proposed a dynamic threshold selection method in 1979. This method suggests maximizing the weighted sum of between-class variances of foreground and background pixels to establish an optimum threshold. We can partition the image into two classes W_1 and W_2 at gray T such that $W_1 = \{0, 1, 2, \dots, T\}$ and $W_2 = \{T + 1, T + 2, \dots, L - 1\}$, where L is the total number of the gray levels of the

image. Let the number of pixels at i gray level be n_i , and $N = \sum_{i=0}^{L-1} n_i$ be the total number of pixels in a given image. The probability of occurrence of gray level i is defined as $p_i = \frac{n_i}{N}$, $p_i \geq 0$, $\sum_{i=0}^{L-1} p_i = 1$. W_1 and W_2 are normally corresponding to the object of interested and the background, the probabilities of the two classes are

$$P_{w1} = \sum_{i=0}^T p_i \text{ and } P_{w2} = \sum_{i=T+1}^{L-1} p_i = 1 - P_{w1} .$$

The means of the classes W_1 and W_2 can be computed as:

$$\mu_{w1} = \sum_{i=0}^T \frac{i * p_i}{P_{w1}} \tag{6}$$

$$\mu_{w2} = \sum_{i=T+1}^{L-1} \frac{i * p_i}{P_{w2}} \tag{7}$$

we can get the equivalent formula:

$$\sigma^2(T) = P_{w1} P_{w2} (\mu_{w1} - \mu_{w2})^2 \tag{8}$$

The optimal threshold T^* can be obtained by maximizing the between-class variance.

$$T^* = \text{Arg max}_{0 < T < L-1} \sigma^2(T) \tag{9}$$

Genetic algorithm has outstanding advantages of simple, robustness and parallelism to get the optimal result, meanwhile the process of Otsu method finding threshold is to get the optimal result of formula (9), so we can use the improved genetic algorithm to find

the optimal threshold for image segmentation. The following presents the new improved adaptive genetic algorithm combined with Otsu method for image segmentation.

3.1 The Main Steps of Improved Genetic Algorithm Combined with the Otsu Algorithm

- 1) Initialization of population: It is hard to find the optimal result when the size of population is too small or too large. Because on the one hand the search space is limited when the size of population is small, on the another hand the computational complexity will increase when the size of population is large Since the image gray values range is 0 to 255, so we use eight-bit binary encoding to represent the 256 states of the individual in populations.
- 2) Evaluation Criterion: Fitness of individual in populations are calculated by formula (9), the results are generally ascending by order, so that it is convenient to do the processing of next step.
- 3) Selection Operation: We use the gambling roulette selection strategy to finish the selection operation. Proportional selection strategy select a new species in a probability which proportional to the fitness. The populations which have the same size of the original population are selected by Gambling roulette algorithm.
- 4) Crossover Operation : Getting a pair of individuals from the new population, then calculate the crossover probability P_c of them, then to compare P_c with random variable R_1 , if P_c is greater than R_1 , get the start location by random operation to do cross operation, if P_c is little than R_1 , the cross operation of them will be canceled.
- 5) Mutation Operation: Traverse all individuals of the entire population, to calculate the mutation probability P_m , then to compare P_m with random variable R_2 , if P_m is greater than R_2 , get the mutation location by random operation to do mutation operation, if P_m is little than R_2 , the mutation operation of them will be canceled.
- 6) Termination criteria: The algorithm will stop when it reaches the largest algebra or group fitness value still doesn't change after 30 iteration loops, and the fitness value individual will be the optimal threshold value T or turn to step 3.
- 7) Image segmentation: Using optimal threshold value to divide the 256 grey values into two categories. Making all of the grey value of pixels less than T to be 0 and making the gray value of pixels equal or greater than T to 255. This will make the original gray image into a binary image to achieve the parting off between the target image and the background image.

4 The Results of the Simulation Experiment and Analysis

In order to validate the effectiveness of the proposed method, we test the improved genetic algorithm and image segmentation, setting k_1 and k_2 0.3. The results of test show that the genetic algorithm proposed in this paper is more accurate and consume less time, which have better segmentation effect compare with the other two genetic algorithms.

4.1 Algorithm Verification

In order to confirm the robustness and parallelism of the proposed method, we compare it with the algorithm [12] and [13]. Three algorithms all adopted the binary encoding and classic gambling roulette selection algorithm, the length of individual is 12 binary, the size of population is 40,the maximum genetic number is 150,and based on the characteristics of the single-variable multi-peak of Otsu method, we test the three Genetic algorithms 50 times by the functions F1 and F2.

1) Function F1:

$$\max F1 = x * \sin(10\pi x) + 2.0, -1 < x < 2$$

This function has the maximum value of global 3.8503 in the range (-1, 2).

2) Function F2:

$$\max F2 = x^2 - 10 * \cos(2\pi x) + 10, -5.12 < x < 5.12$$

This function has the maximum value of global 40.3533 in the range (-5.12, 5.12).

The optimal solutions comparison and evolution number of three genetic algorithms for function F1 show in Fig.1 and Fig.2. The optimal solutions comparison and evolution number of three genetic algorithms for function F2 show in Fig.3 and Fig.4.

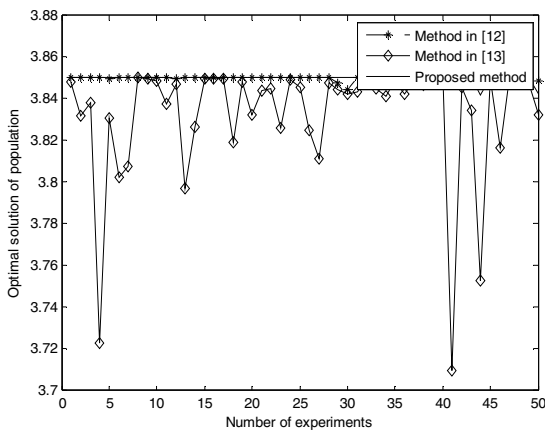


Fig. 1. The comparison of optimal solutions for function F1

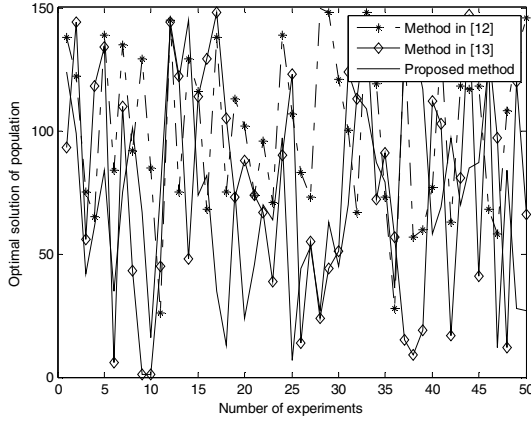


Fig. 2. The comparison of evolution number for function F1

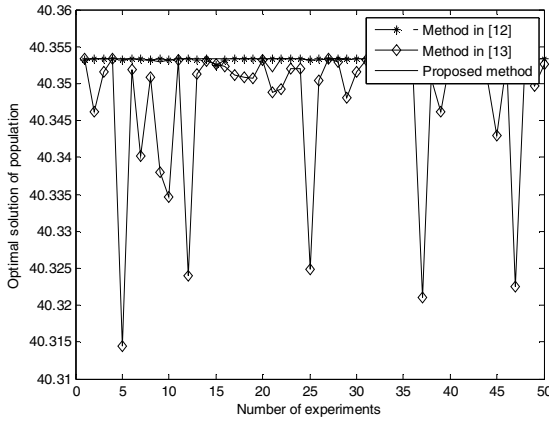


Fig. 3. The comparison of optimal solutions for function F2

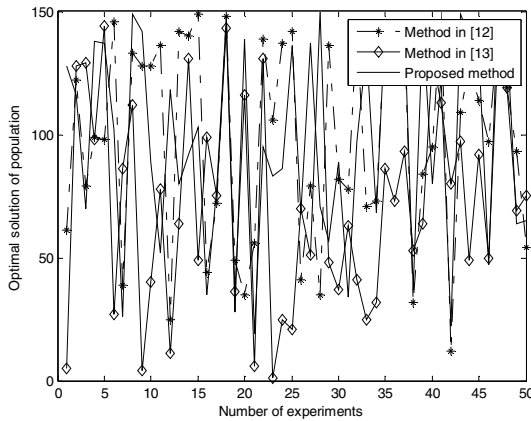


Fig. 4. The comparison of evolution number for function F2

We can get following conclusions from table 1. The method [12] could converge quickly to a solution, but it isn't stable and accurate; the method [13] could provide more accurate and stable solution, but the rate of getting optimal solution is not as fast as the former; on the other hand the proposed method could find the optimal solution as fast as the method [12], meanwhile could provide more stability and accuracy then the method [13].

Table 1. Test results of three kinds of adaptive genetic algorithm

		Method [12]	Method[13]	Proposed method
F1	Optimal solution	3.8503	3.8503	3.8503
	Average solution	3.8379	3.8488	3.8502
	Average number of evolution	62.5600	105.7000	65.2000
	The variance of optimal solution	6.6908e-004	8.8319e-006	1.0934e-008
F2	Optimal solution	40.3533	40.3533	40.3533
	Average solution	40.3436	40.3533	40.3533
	Average number of evolution	67.5800	96.4400	70.0200
	The variance of optimal solution	4.8576e-4	1.0205e-9	3.1215e-10

4.2 The Effect of Image Segmentation

In order to evaluate the performance of the proposed method, our algorithm has been tested using images 1 (256×256) and image 2(256×170) in Fig.5. We can see from Fig. 6 that the proposed method produces images that are successfully distinguished from the backgrounds.



(a)



(b)

Fig. 5. Original images: (a) image 1, (b) image 2



Fig. 6. Segmentation results: (a) image 1, (b) image 2

We compare the proposed method with the method [12] and [13] by repeat the test 20 times. The test results showed in table 2. The proposed method and the method [12] and [13] are based on Otsu method which finds the optimal threshold by brute force; so that we can believe that the optimal threshold of Otsu method is most accurate and as the standard threshold.

Table 2. The run time and thresholds

		Otsu method	Method in [12]	Method in [13]	Proposed method
Image 1	Average optimal threshold	118.20	117.54	118.25	118.19
	Average time/ms	28.430	8.671	10.161	9.015
Image 2	Average optimal threshold	138.15	138.39	138.19	138.15
	Average time /ms	20.891	6.837	8.261	7.024

We can get following conclusions from table 2. The time of Otsu method getting optimal threshold is longest and the method [12] consume the least time, but it's optimal threshold is more different with the standard threshold than the other two method. The optimal threshold by method [13] is more close to the standard threshold, but the time is longer than the method [12]. Compared with the method [12] and [13], the proposed method would get the optimal threshold as fast as method [12], meanwhile the optimal threshold is more close to the standard value than method [13].

5 Conclusions

The method proposed in this paper adjust the parameters of genetic algorithm adaptively according to the difference between the individuals and the decentralizing degree of individuals of the population, so that to optimize the adaptive genetic algorithm. Simulation experiments have proved that the thresholds is more accurate and it consumes less time greatly by using the proposed method, compared with Otsu image segmentation and other genetic algorithm based image segmentation. We

believe that the adaptive genetic algorithm will commonly applied in the image segmentation as the theory of adaptive genetic algorithm further perfected and improved.

Acknowledgment. This work is supported by the National Natural Science Foundation of China under Grant No. 71161007, the Social Science Fund Project of Ministry of Education under Grant No. 10YJCZH049, the Key Science and Technology Program of Haikou under Grant No. 2010-0067 and the Scientific Research Initiation Fund Project of Hainan University under Grant No. kyqd1042.

References

1. He, J., Ge, H., Wang, Y.-F.: Survey on the methods of image segmentation research. *Computer Engineering & Science* 31(12) (2009)
2. Xiao, R.: An image segmentation algorithm using genetic strategy. In: 2010 2nd International Conference Computer Engineering and Technology (ICCET), vol. V6, pp. 605–607 (June 2010)
3. Kang, W.-X., Yang, Q.-Q., Liang, R.-P.: The comparative research on image segmentation algorithms. In: 2009 First International Workshop on Education Technology and Computer Science, Wuhan, Hubei, China, pp. 703–707 (2009)
4. Dong, J.P., Ling, N.: The basic ideas of genetic algorithm segmentation. In: ISCAS 2008, pp. 628–631 (May 2008)
5. Montazersadgh, F.H., Fatemi, A.: Study on genetic algorithm, SAE Paper No. PFL34, Society of Automotive Engineers (2007)
6. Xu, L.: Image Segmentation Based on Improved Genetic Algorithm. In: 2008 International Symposium on Distributed Computing and Applications for Business Engineering and Science, Dalian, pp. 269–273 (2008)
7. Qiao, W.-W., Wu, C.-M.: 2-D thresholding segmentation method based on maximum inter-class cross entropy. *Journal of Northwest University (Natural Science Edition)* 38(3), 374–378 (2008)
8. He, C.-H., Hu, Y.-C.: Automatic threshold segmentation approach based on improved genetic algorithm. *Computer simulation* 28(2), 312–315 (2011)
9. Zhong, Q.U.: Research on Image Segmentation Based on the Improved Otsu Algorithm. *Computer Science* 36(5), 276–278 (2009)
10. Wang, H.-Z., Dong, Y.: An improved image segmentation algorithm based on Otsu method. In: Proceedings of the SPIE, vol. 6625, pp. 1–8 (2007)
11. Srinivas, M., Patnaik, L.M.: Adaptive probabilities of crossover and mutation in genetic algorithm. *IEEE Trans Systems Man and Cybernetics* 24(4), 656–667 (1994)
12. Li, Q., He, W.H.: A study on image segmentation by an improved adaptive algorithm. In: Proceedings of the 6th International Conference on Machine Learning and Cybernetics, Hong Kong, pp. 1570–1573 (2007)
13. Li, K.-S., Li, M.M., Zhang, W.-S.: New method of image segmentation based on improved genetic algorithm. *Application Research of Computers* 26(11), 4364–4367 (2009)

Resistance Identifier Recognition Based on Wavelet Transform and LBP Operator

Chong-quan Zhong and Yan-dong Zhu

School of electronics and information engineering, Dalian Univ. of Technology,
LiaoNing Dalian 116024, China
zhongcq@dlut.edu.cn

Abstract. In order to promote the accuracy and practicability of resistance identifier feature recognition in the defects inspection of PCB (Printed Circuit Board), a resistance texture feature description and recognition method combining wavelet transform and Ojala's LBP (Local Binary Pattern) operator is proposed in this paper. Firstly, wavelet analysis is adopted to decompose the original resistance image for dimension reduction. Then the approximate image is divided into several sub-blocks, and two types of sub-block LBP histogram is drawn by using two different LBP operators. Finally, we concatenated the whole histogram of every sub-blocks into histogram sequence, and the sequence is the enhanced feature vector of resistance image recognition. Experimental results show that the proposed method has a high-level recognition rate of texture feature.

Keywords: Wavelet Transform, LBP operator, Normalizing Histogram, Texture Feature, Resistance Recognition.

1 Introduction

In electronic industry, PCB is the main part in many electronic applications. The defects of PCB may lead to scrape, and even more, the whole product. So the detecting and testing of PCB is an indispensable link for quality control. Chip resistance plays an important role of the onboard component, its extraction and identification of identifier feature is a key for PCB's defects inspecting.

Because the appearance and size of resistance is the same, the only way to distinguish the difference is to inspect and recognize the characters of resistance's surface. Recently, the methods of the resistance recognition can be divided into two types: Character Recognition, and Texture Analysis[1]. Because there is not a unified standardized character and word size for the components, it's hard to keep the quality of Character Recognition satisfied by using projection feature[2], space distribution[3]and shape context methods. However, the Texture Analysis methods depend on the resistance's labeling, which is easy to achieve the inspection and recognition requirement of the resistance. Based on the basic texture analysis method, we found many advanced methods of it, which can be divided into two types: statistical method and structure method. The advantage of the statistical method lies in its simple principle and easy to realize, while the disadvantage is its limited application range. Thus,

the structure method is more suitable for texture primitive analysis and arrangement rules.

The paper puts forward an improved method of a resistance texture analysis. As an application of wavelet decomposition, it is highly effective for the removal of redundancy information, the dimensionality reduction of image and the calculation time reducing. By using wavelet decomposition, we obtained an approximate image of the resistance, and divided it to sub-blocks. Then we calculated the LBP eigenvalue of each sub-block by using two different kinds of LBP operator, and draw the feature histogram. Finally, we concatenated the whole histogram of every sub-blocks into histogram sequence, and achieved the resistance’s discriminant feature. As a result, the proposed method is a combination of wavelet analysis and LBP operator; it can obtain more abundant and effective discriminant features, and increase the expression ability and distinguished ability of the spatial histogram. Thus the method can increase recognition rate effectively and show significant practical application value.

2 The Basic Characteristics Operator

For the chip resistance, the characters of its surface are the main texture feature. LBP is an efficient operator because of its excellent description capability of local texture. The use of LBP[4] operators make it easy to extract the texture features of image.

2.1 LBP Operator

The basic LBP operator is a fixed size 3×3 rectangular pixels block $f(x_c, y_c)$, The rectangle has a central pixel g_c and eight neighboring pixels g_0, \dots, g_7 , so there are nine gray levels. The basic LBP operator can be found in Figure 1. First, we compare the gray level between central pixel and other eight neighboring pixels. If the gray level of neighboring pixel is larger than the central pixel, then the value of the pixel is 1; otherwise the value is 0. Along the clockwise direction, it will be calculated to eight binary values, which is the eigenvalue of the 3×3 rectangular pixels block. By scanning the whole area, a histogram can be drawn, which consists of all the eigenvalues in the area. And the histogram is the texture description in the area.

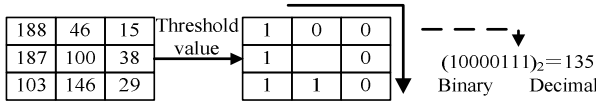


Fig. 1. The basic LBP operator

Define T as the joint distribution of gray levels, and $P+1(P>0)$ is the number of gray levels, and the equation of texture is shown below:

$$T = t(g_c, g_0, \dots, g_{P-1}) \tag{1}$$

T: joint distribution of gray levels (texture), P: number of gray levels; t: function of joint distribution, g_c : gray level of central pixel, g_p ($p=0, \dots, \tilde{P}-1$): gray level of local neighborhood pixels (equal interval to the central pixel).

If there is no information lost, gray level of central point is threshold value; and the binary value of other pixels is calculated.

$$T = t(g_c, g_0 - g_c, \dots, g_{P-1} - g_c) \tag{2}$$

Because of the reducing and enlargement of the image size and scale, the gray level difference can be affected. Therefore an equation $s(x)$ is defined to avoid it:

$$s(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \tag{3}$$

From equation(2) and equation(3), it can be written as:

$$T \approx t(s(g_0 - g_c), \dots, s(g_{P-1} - g_c)) \tag{4}$$

Thus, the LBP value of the window, which is an 8-bit binary value, is shown below:

$$LBP(x, y) = \sum_{p=0}^7 s(g_p - g_c) 2^p \tag{5}$$

In order to adapt to different scales of texture features, Ojala extends LBP operator of the 3×3 rectangular neighborhood to any circle neighborhood. The method of bilinear interpolation can obtain the gray level of the point not completely fall on the pixel position. In addition, the improved LBP operator allows the radius R of the circular neighborhood has an arbitrary number of pixels. The extended LBP operator is shown below in figure 2.

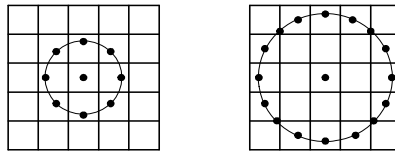


Fig. 2. The extended LBP operator: the circular(8,1) and the circular(16,2) neighborhoods

As the basic LBP operator is limited for the description of texture features, Ojala proposed an extended LBP operator – uniform mode. When the eigenvalue of LBP operator is considered to the ring linked end to end, there is at most two bit transition from 0 to 1 or 1 to 0; and this LBP operator is called the uniform pattern, such as 00111110. The method greatly reduces the types of binary mode, and the experiment results are very satisfactory. Using the scale (8, 1) of the LBP operator, we have the number of the uniform pattern is 59, which is account for only 23% of the total

number (256), and the percentage of the matched textures is 87.2%[5]. As a result, uniform pattern can effectively describe the majority of the image texture features, significantly reduce the number of features, and contain most of the image texture features. The LBP feature images of different operator are shown in figure 3:



Fig. 3. The LBP feature images of different operator (the original image, the basic LBP operator, and the uniform pattern)

2.2 Discrete Wavelet Transform

In the process of resistance image acquisition, we need to reduce the noise for keeping the image quality.

The noise of resistance image is mainly distributed on high-frequency components. Therefore, there is a high-frequency coefficient of wavelet for denoise and compression. It can reduce the amount of data, and also retain the most effective information.

Wavelet decomposition of digital images is the two-dimensional discrete wavelet transform. In the transformation of each layer, the image is decomposed into four 1 / 4 size of it. Two-dimensional image data as discrete data is given, with $f(x,y)$. First, the horizontal direction discrete wavelet transform is preferred, then the vertical direction.

By using Harr wavelet decomposition of image data, set $L_2(R^2)$ to meet the two-dimensional image data gives $f(x,y)$, $x,y \in Z$, set the scale factor as 0, $C_{0,x,y} = f(x,y)$ is given, and the two-dimensional image data of the wavelet decomposition formula is:

$$\begin{aligned}
 C_{j,x,y} &= 2 \sum_{k,l} h_{k-2x} h_{l-2y} C_{j-1,k,l} \\
 d^1_{j,x,y} &= 2 \sum_{k,l} h_{k-2x} g_{l-2y} C_{j-1,k,l} \\
 d^2_{j,x,y} &= 2 \sum_{k,l} g_{k-2x} h_{l-2y} C_{j-1,k,l} \\
 d^3_{j,x,y} &= 2 \sum_{k,l} g_{k-2x} g_{l-2y} C_{j-1,k,l}
 \end{aligned} \tag{6}$$

Where $\sum_k h_k = 1$, $g_k = (-1)^n h_{1-n}$, $\sum_k g_k = 0$, $C_{j,x,y}$ is approximation coefficient LL_k , $d^1_{j,x,y}$ is level detail coefficients HL_k , $d^2_{j,x,y}$ is vertical detail coefficients LH_k , $d^3_{j,x,y}$ is diagonal detail coefficients HH_k , Mallat 2-Level wavelet decomposition is shown below in figure 4.

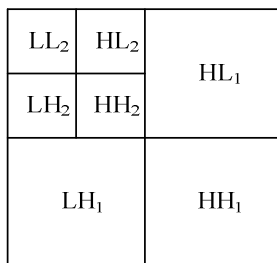


Fig. 4. Mallat 2-Level wavelet decomposition

First, we translate the true color image of resistance to the grayscale space, and then decompose the resistance image by using the method of two-dimensional discrete wavelet transform. We set the image function of input resistance to $f(x, y)$, 2-D wavelet decomposition for image is shown below in figure 5:

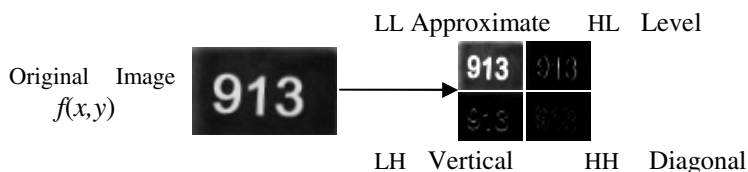


Fig. 5. 2-D wavelet decomposition for image

These components have the same size as the amount of data; but the original image information is different or even irrelevant. The low-frequency (approximation) data of wavelet decomposition contains most of the original image information. The high-frequency (detail) components of the wavelet decomposition express a different direction details, such as the horizontal, vertical and diagonal direction.

For texture feature description and recognition, we need the effective information of the image itself. A large amount of information through layers of wavelet decomposition should be excluded. So based on real-time and processing speed requirements of system detecting, the only way is using the low-frequency (approximation) images for two-dimensional discrete wavelet decomposition.

3 Improved Local Binary Pattern (LBP) Features

How to effectively extract the feature is the key of resistance classification and recognition. LBP features have good local properties, strong gray-scale invariance and rotation invariance. But when we use LBP operator for feature extraction, it also has the following problems: (1) Local binary pattern cannot be a good description of the spatial structure image information, that is the global information.(2) When calculating the LBP operator by using the pixel gray value comparison, if the LBP operator of the original image is used directly, the noise will be large. (3) The LBP operator can only

be used to extract the local texture information, and one LBP operator can only extract one certain scale of feature information. Thus, the feature extraction capability is poor.

For the latter two issues, we proposed a wavelet decomposition of the image firstly, and then divide the approach image into sub-blocks. Last, we used features histogram of wavelet division by different LBP operators to solve the problems.

3.1 Algorithmic Process

In order to obtain the sub-block feature histogram of LBP operator, we put forward the following steps: (1) We selected the appropriate wavelet for wavelet decomposition, and then use an approximation image of 2-D wavelet decomposition instead of the original one, that is the filtering process. This process can effectively overcome the noise, reduce dimensionality of the down-sampling process, and greatly save the calculation time. (2) We divided the approximation image into multiple non-overlapping sub-blocks, which have their specific size. (3) Two types of sub-block LBP histogram is drawn by using two different LBP operators. (4) Finally, we concatenated the whole histogram of every sub-blocks into histogram sequence, and the sequence is the enhanced feature vector of resistance image recognition. The flowchart of histogram feature is shown below in figure 6:

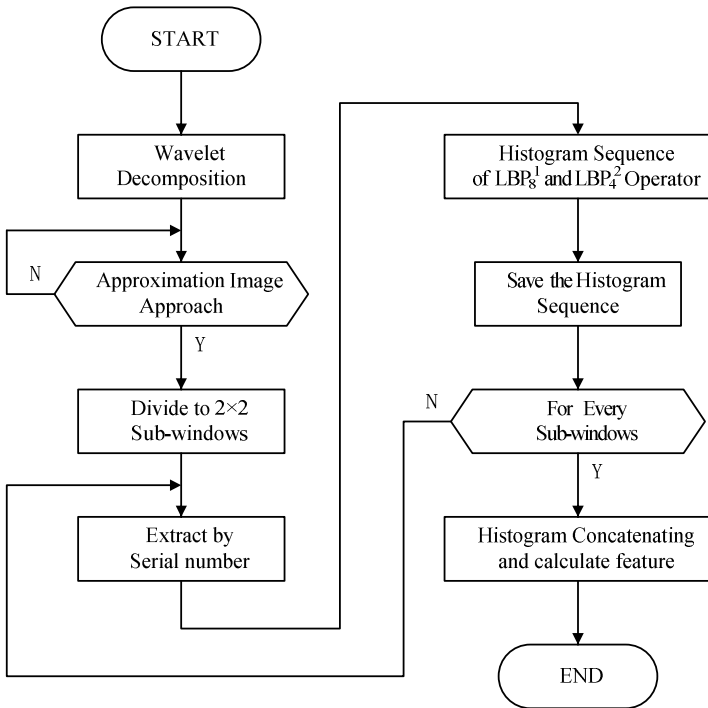


Fig. 6. The flowchart of histogram feature by using wavelet division local binary patterns method

3.2 Feature Extraction and Detection

Standard LBP operator can extract a lot of features, but some did not provide good texture information. We need to divide the approximation image of the wavelet decomposition into sub-windows. From combining the LBP features of every sub-window, we draw a one-layer histogram. Sub-division of the window use window stacks strategy[6].

First the approximation image is decomposed into 2×2 sub-windows. Then we calculated LBP features to represent local information of the image, the calculation of the histogram is revised to:

$$H_{i,j} = \sum_{x,y} I\{f_1(x,y) = i\} \quad I\{(x,y) \in R_j\}, i = 0, \dots, n-1; j = 0, \dots, m-1 \quad (7)$$

R_0, R_1, \dots, R_{m-1} : represent each one of these sub-blocks; i : said a LBP pattern, which depends on the LBP operator

For the LBP_8^1 operator, it has eight samples, the range of which is [1, 59]; and for the LBP_4^2 operator, it has four samples, the range of which is [1, 15].

I is defined as:

$$I(A) = \begin{cases} 1, & A > 0 \\ 0, & A \leq 0 \end{cases} \quad (8)$$

To describe the standard information of the resistance image, we need to get the global features. In order to get the features of whole image, we concatenated the histogram of every sub-block into histogram sequence using each LBP operator. and used the sequence as the feature vector of resistance image. m -block area of the histogram is expressed as $H_m, m=1, 2 \dots n \times n$, and $n \times n$ means the image is divided into $n \times n$ sub-blocks. For the two kinds of LBP operator uses different blocking method. The histogram features of whole image can be expressed as:

$$H = \begin{cases} (H_1, H_2, \dots, H_{n1 \times n1})_1 \\ (H_1, H_2, \dots, H_{n2 \times n2})_2 \end{cases} \quad (9)$$

$(H_1, H_2, \dots, H_{n1 \times n1})_1$: The histogram sequence using the LBP_8^1 operator

$(H_1, H_2, \dots, H_{n2 \times n2})_2$: The histogram sequence using the LBP_4^2 operator

All these two methods divided the image into sub-blocks, and the scale of the LBP feature is $59 \times 4 + 15 \times 4 = 296$, that is, the feature vector length of the image is 296. The result is shown below in figure 7:



Fig. 7. The flowchart of the method

3.3 Similarity Measure

With the log-likelihood statistical methods to measure their similarities and differences, we calculated as follows:

$$m = \sum_{i=1}^n \left| \lg\left(\frac{H_1[i]}{H_2[i]}\right) \right| \tag{10}$$

H_1 : The improved LBP histogram image of the standard resistance

H_2 : The improved LBP histogram image of the testing resistance

For the result m , we select the appropriate threshold. If the calculated value of equation (10) is greater than the threshold value, the detected resistance has defects. Many experimental studies show that the threshold is taken as 10.

4 Analysis and Application of the Experimental Results

A set of experimental images, as is shown in figure 8. Where (a) is the standard image, and the remains are test images. LBP features were calculated for each image. According to equation (10), the results of image (b) ~ (f) as follows: (b)5.4 (c)19.9 (d)6.3 (e)19.7 (f)17.2. Because the value of (b) and (d) is smaller than the threshold, it is the same kind of resistance to the image (a). But the image of (c), (e), (f) are different.

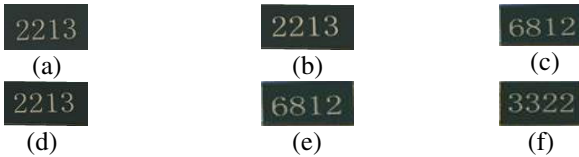


Fig. 8. Experiments of resistances test

To evaluate the performance of the proposed method in this paper, we built a resistance image database for the algorithm comparing and testing. The database is composed of five different resistances, for each resistance, it has 10 images corresponding to a different light and location. Comparing with the three different methods, which are character recognition method, basic LBP window method and the improved LBP window method. We work out the experimental results showing that the optimal one is the proposed method of the improved LBP window, which has a high-level recognition rate of texture feature. The comparison result is shown below in table 1:

Table 1. Comparison of different methods

method	recognition rate
Improved LBP window	0.91
Character recognition	0.82
Basic LBP window	0.85

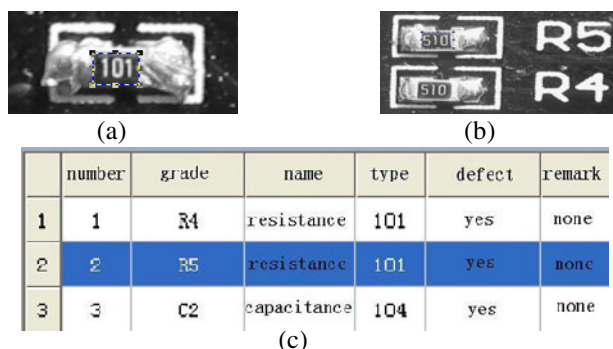


Fig. 9. Display of the detection steps and the result

In the system of PCB defects detection, we set resistance testing window as figure 9(a) shown firstly. Then we chose the option of welding feature detection by right click, and it is shown in figure 9(b). The program of standard image defects inspection will store the feature information to the database. This information corresponds to the feature of the test point and the appropriate location of the window. By using equation (10), we considered the welding is normally when the threshold is satisfied. Otherwise, the defect is considered detected. At last, we store the result record to the database which is a control of GridCtrl. The results are shown in Figure 9(c).

5 Conclusion

Resistance defects detection is the key of PCB detection. The precision and accuracy of detection will directly affect the system performance and efficiency. Based on the method of wavelet transform, an improved resistance texture feature description and recognition method is proposed. The method achieved the dynamic division of resistance images in the variable-size sub-windows. And by using LBP operators of different uniform patterns, we can extract the features of the sub-window easily. The experimental results show that the improved resistance texture feature description and recognition method is much more effective for the defects detection of resistance.

References

1. Liu, M.: Automatic Defects Inspection System of Printed Circuit Board Based on Vision. DaLian University of technology, DaLian (2006)
2. Zhao, C., Shi, Y.: A Method of Quickly Recognizing Chinese Character in Vehicle Plate Based on Projection Character. *Computer Engineering and Applications* 41(19), 207–209 (2005)
3. Wu, J., Qiu, Y.: Handwritten Numeral Recognition Based on Space-distribution Characteristics. *Journal of Wuhan Yejin University of Science and Technology* 27(2), 176–178 (2004)

4. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition* 29(1), 51–59 (1996)
5. Zhou, Y., Zhou, Z., Sun, N.: Leukocyte Classification Based On an Enhanced Local Binary Pattern Feature. *Journal of Biomedical Engineering Research* 4(16), 242–246 (2006)
6. He, L., Zou, C., Zhao, L.: An enhanced LBP feature based on facial expression recognition. In: *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China* (2005)
7. Pietikäinen, M.: Image Analysis with Local Binary Patterns. In: Kalviainen, H., Parkkinen, J., Kaarna, A. (eds.) *SCIA 2005*. LNCS, vol. 3540, pp. 115–118. Springer, Heidelberg (2005)
8. Mao, B.: Texture analysis based on adaptive LBP. *Computer Engineering and Applications* 43(2), 89–90 (2007)
9. Zhang, D.: An Image Retrieval Algorithm Researching Based on Color Feature and Local Binary Pattern. *Microcomputer Applications* 30(6), 35–38 (2009)

Image Processing Methods for V-Shape Weld Seam Based on Laser Structured Light

Tao Qin, Ke Zhang, Jingyu Deng, and Xin Jin

Shanghai Key Laboratory of Materials Laser Processing and Modification,
Shanghai Jiaotong University, Shanghai 200240, China
{jetqin,zhangke,jinxin}@sjtu.edu.cn, dengjingyu86@163.com

Abstract. Fast and reliable image processing method is one of the key technologies for realizing welding seam tracking. To this end, a new image processing method for V-shape weld seam based on laser structured light is proposed. This method starts with image preprocessing, including median filtering, one-dimensional LOG (Laplace of Gaussian) operations, morphological operations, and scattered clusters filtering. For improving the efficiency of image processing, it can find a target window dynamically. Then, the local search method is used twice for improving the accuracy of extracting the laser centerline. At last, this laser centerline is fitted by the modified least square method, and the feature points of V-shape weld seam can be obtained by calculating the intersections of the fitted lines. The results of image processing show that this method can quickly and reliably process seam images with strong interference, and meet the requirements of welding seam tracking.

Keywords: Image processing, Dynamic window, Centerline, Least square method.

1 Introduction

With the development of machine vision technology, seam tracking methods which are based on the visual sensor have become hot topics in the field of intelligent welding [1-4]. According to the types of visual sensor, these methods can be generally divided into active vision methods and passive vision methods. However, passive vision methods which are based on laser structured light, because of its wide adaptability, good stability, high accuracy, and proper anti-jamming, become one group of the major seam tracking methods. Although passive vision methods have the above mentioned advantages, due to the arc, the smoke and the splash under the real condition of weld, the quality of images captured by visual sensor is uneven. In order to guarantee the reliability of the seam tracking system, fast and reliable image processing methods are particularly important. In this paper, a new image processing method for V-shape weld seam is proposed. It mainly improves the dynamic window method, extracting centerline method, and fitting line method. The following sections in this paper describe this method step by step and use an weld image with proper interferences to verify the effect.

2 Image Preprocessing

Good image preprocessing method can reduce the complexity of the extraction, and improve the accuracy of the extraction result. The details of the image preprocessing method are described in section 2. The weld image acquired by the CCD camera is 8 bit gray image, with the size of 656 by 494 pixels, as shown in Fig.1. The origin of the image is the upper left point, The X axis is horizontal and the Y axis is vertical.

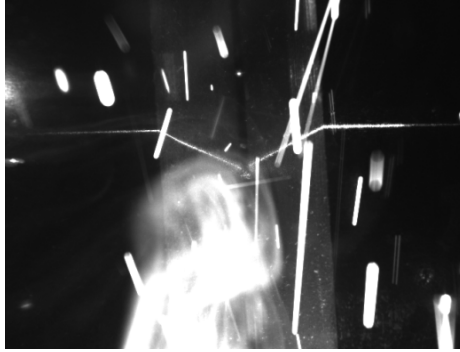


Fig. 1. The weld image

2.1 Median Filtering

The aim of Median filtering is to eliminate the noises, produced during the process of image acquisition, quantization, transmission, which have the characteristics of randomness and discreteness. This paper chooses 3 by 3 rectangular median filter to process the weld image, for showing the effect of the median filtering, partial enlarges of the center regions in the image before and after processing are shown in Fig.2.

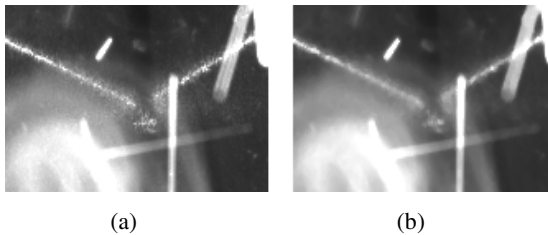


Fig. 2. Partial enlarges of the center regions: (a) Before filtering; (b) After filtering

2.2 LOG Filtering

The width in the vertical direction of the laser stripe is very narrow and nearly a constant, however, the width of the arc, splash, and smoke in the same direction far outweigh that of the laser stripe. LOG filtering can reduce these large interferences and

enhance the laser stripe [5]. In this paper, the filter expressed in expression (1) is used as the width of the laser stripe is about 5 pixels.

$$T = [-2, -2, -1, 1, 2, 4, 2, 1, -1, -2, -2]^T. \quad (1)$$

The result of LOG filtering is shown in Fig.3.

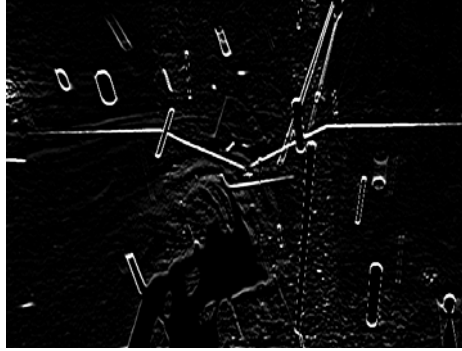


Fig. 3. Image after LOG filtering

It can be seen in Fig. 3 that large blocks of interferences are hollowed out, leaving only edges. Furthermore, laser stripe and interferences having the similar width as the former are enhanced. Although these interferences are enhanced, because of their isolated distribution, they can also be eliminated by the follow-up Particle filtering.

2.3 Self-adapting Dynamic Window

As the distribution of the laser stripe is level, the maximum value accumulated by pixel value in each row is to be gotten around the centerline of laser stripe. To find the row index, denoted as Y_H , firstly, creating a matrix, denoted as I , with the size of 656 by 494 and each element is the pixel value which has the same row and column indexes in the image. Secondly, using a column vector, denoted as c , with the size of 656 and all elements are one, to post-multiply the matrix I , and produce a column vector, denoted as d . Thirdly, finding the index Y_H corresponding to the maximum element in d . And the window area can be decided by expression (2).

$$\begin{cases} Y_{\max} = \min\{Y_H + h + \Delta, 655\}; \\ Y_{\min} = \max\{Y_H - \Delta, 0\}. \end{cases} \quad (2)$$

where Y_{\min}, Y_{\max} are the upper and lower row indexes, h equals to the approximation value of the depth of weld groove, Δ equals to width of edges in the window. In this paper, h is set to 80 and Δ is set to 40.

The above method can also realize the window method proposed in literature [6]. However, the distribution of the laser stripe cannot always be level, so the accurate dynamic window couldn't be guaranteed.

To get more accurate results, the above method can only be used at the initial moment in the welding process when the distribution of the laser strip is level. As the time interval between the adjacent weld images is very short (the fps of the CCD camera is 30), so the displacement of the laser stripes between the adjacent images is very small, based on this fact, the above method is modified as follows. According to the image processing results of the previous image, three intersections can be calculated (see section 3, 4). $Y_{n-1,l}$ is Y coordinate of A1, which is the intersection of the left level part of the laser stripe and the left border of the previous image, $Y_{n-1,v}$ is Y coordinate of A2, which is the intersection of the two slant parts of the laser stripe, $Y_{n-1,r}$ is Y coordinate of A3, which is the intersection of the right level part of the laser stripe and the right border of the previous image. The relative position of the three intersections is described in Fig.4.



Fig. 4. Schemes of positions of the intersections

The range of the window is determined by the maximum and minimum value among the three intersections, as is shown in expression (3).

$$\begin{cases} Y_{n,max} = \min\{\max\{Y_{n-1,l}, Y_{n-1,v}, Y_{n-1,r}\} + \Delta, 655\}; \\ Y_{n,min} = \max\{\min\{Y_{n-1,l}, Y_{n-1,v}, Y_{n-1,r}\} - \Delta, 0\}. \end{cases} \quad (3)$$

where $Y_{n,max}$ and $Y_{n,min}$ are the upper and lower row indexes of the present window. Δ has the same meaning as that in expression (2).

Compared to the former window method, the latter one make good use of the information of the previous image, consequently it can find window first, then process the window. Fig.5 is the window in the Fig.3 that is processed by the latter method. The size of the window is 656 by 160 pixels.

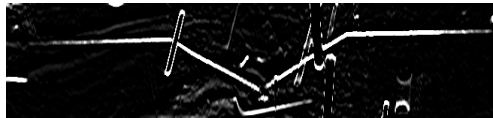


Fig. 5. Target window in Fig.3

2.4 Threshold Segmentation

There are many threshold segmentation algorithms, such as manual thresholding, maximum entropy thresholding, Otsu thresholding and iterative thresholding. The threshold value is calculated from the histogram of Fig.5 (see Fig.6).

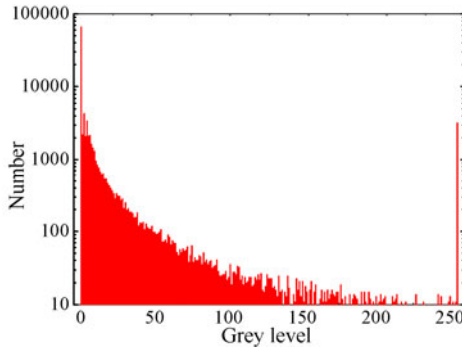


Fig. 6. Histogram of Fig.5

From this histogram, it can be seen that the pixels whose value is 0 or very small account for the majority of all the pixels in the Fig.5. On the contrary, the number of the pixels which have the value above 200 is very small. Considering the efficiency and adaptability of the algorithm, iterative thresholding is used to process the Fig.5 and the initial value is set to 128. The result is shown in Fig.7. It consists of two pixel value: 0 and 255, representing the black and white area respectively.



Fig. 7. Image after iterative thresholding

2.5 Particle Filtering

Although the strong interferences, such as the large chunks of splash, smoke, have been hollowed out by LOG filtering, their edges still exist. To remove this kind of edges and other isolated particles, the particle filtering should consist of morphological operation and marking algorithm. Specifically:

(1) Use morphological open operation to process Fig.7. The result is denoted as Fig. A.

(2) Use marking algorithm which is based on eight-connected domain to mark the white connected particles in Fig. A. Set a threshold for the number of the pixels consisted in these particles, and remove the particles below the threshold, and keep the particles above the threshold. The result of particle filtering is shown in Fig.8.



Fig. 8. Image after Particle filtering

3 Extracting the Laser Stripe Centerline

In literature [7], centerline is extracted by morphological thinning operation, but at the same time, much burr is produced, and that is not beneficial to the fitting line processing. In literatures [8], [9], the edges of the laser stripe are extracted first, and then the centerline is extracted by calculating the mean value of indexes of the upper and lower edges. This method is simple and easy to program. In this paper, we improve this method, and neglect the extraction of the laser stripe. We call this extracting method Double local search method. To describe our method clearly, (x, y) is denoted as coordinates of pixels in the image, *Error* can be any real number bigger than width of the window along the Y direction, and it marks the invalid points in the centerline, w presents the width of laser stripe in vertical direction.

Considering search from the left border to the right border of the window, Double local search method is described step by step as follows.

(1) In the first column of the window, search the pixel whose value is 255 from the first row in the downward direction and the last row in the upward direction respectively. When the first pixel whose value is 255 is found in the both directions, note down the Y coordinates of the two pixels as y_1, y_2 . If these pixels cannot be found, denote $(x, Error)$ as the coordinates of the point in the centerline in this column, and jump to step (3).

(2) Calculate the difference between the two y noted down in the first step, denoted as Δy . If Δy is not larger than w' , which could be a little larger than w , then calculate the mean value of the two y , denote $(x, (y_1+y_2)/2)$ as coordinates of the coordinates of the point in the centerline in this column. And else, denote $(x, Error)$ as the coordinates of the point in the centerline in this column.

(3) x plus 1, if x is larger than 655, then jump to step (5). And else if the previous y doesn't equal to *Error*, the search range in the current column can be narrowed, which is $[y-w, y+w]$. If y equals to *Error*, then the search range is still from the first row to the last row of the window. The search method is the same as that proposed in step (1).

(4) Repeat step (2) and (3).

(5) End.

Then use the algorithm to extract the centerline from the right border to the left border of the window. Searching twice can produce two centerlines, in most cases, they are not the same. After observing large number of weld images, the interferences, like arc and splash, in the upper region of the image is much weaker than that in the lower region which is closer to the welding pool. Based on this judgment, we compare the two centerlines, note down the smaller y as the Y coordinate in columns and remove the invalid points. This centerline could consist of short lines, long lines, points, and maybe short curves. So at last, based on the connectivity, the centerline can be refined by removing its short lines, points and short curves. For clear description of the centerline, change the black background to grey, the white centerline to black. The result is shown in Fig.9.



Fig. 9. Extraction of laser stripe centerline

4 Fitting Line

Hough Transform is widely used to extract straight lines, but the Standard Hough Transform has many disadvantages such as poor precision, requiring large memory, and time-consuming [10-12]. This paper proposes the modified least square method which can cover the disadvantages of the Standard Hough Transform.

The least square method cannot fit a proper line in a image consist of several lines and isolated points which are randomly distributed, that is due to its sensitivity to the noise and lack of discrimination on the optimal line. In this paper, we define the optimal line as the line that contains the most points to be fitted and meets certain linearity requirement. For reducing the sensitivity and adding the ability of discrimination, the following parameters should be defined.

(1) The constant R . We assume that a point belongs to a line if the point lies within the distance R from the fitted fine.

(2) The mean square value of the distances between the points which are used to fit the line to the fitted line, denoted as M .

(3) The linearity of the fitted line, denoted as l , when l is closer to 1, the fitted line is more perfect.

$$l = 1 - \frac{M}{R^2}. \quad (4)$$

(4) The variable N . We eliminate N farthest points from the fitted line when it needs to be refitted.

The fitting line method is detailed as following steps:

(1) Fit a line from the initial point set and eliminate N farthest points from the fitted line and then refit the remaining point set until the linearity of the fitted line meets linearity requirement. We assume that the line could be the optimal line, and store the points that belong to this fitted line. And then we remove these points from the initial point set and store the remaining point set.

(2) Repeat step (1), the previous remaining point set is to be the initial point set in the next circulation, until almost all the points are rationed among the corresponding fitted lines.

(3) Fit the line from the point set having the most points, and this line can be regarded as the optimal line. If the linearity doesn't meet the requirement. We can also use step (1) to refit the line until the linearity meets requirement.

The program flow chart is shown in Fig.10 according to the above steps.

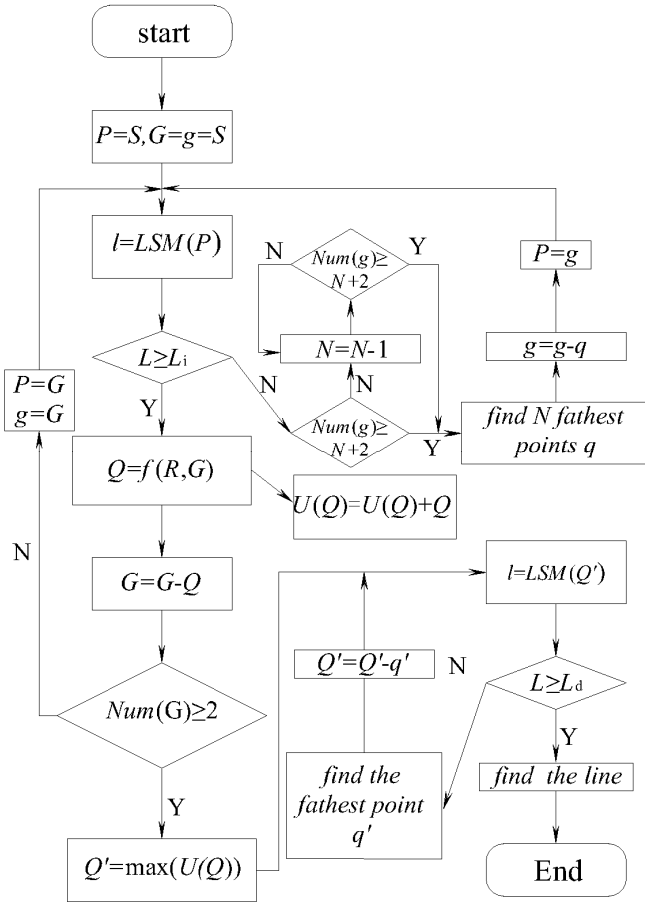


Fig. 10. The program flow chart of the fitting line method

The symbols and expressions in the flow chart are explained as follows:

(1) S represent the initial point set, P, G, g, Q, Q', q , represent corresponding point set used to store the points, q' represents the farthest point from the fitted line in step (3). $U(Q)$ store the point set belong to the corresponding fitted line. L_i is the requirement for linearity in step (2), L_d is the requirement for linearity in step (3), which is a little larger than L_i . To ensure the accuracy, R can be set to 1 or 2, the initial value of N is set to 5 in this paper.

(2) $f(R, G)$ represents the subset of point set G and it belongs to the fitted line, $LSM(P)$ calculates the linearity of the fitted line fitted by applying the least square method to point set P , $\max(U(Q))$ represents the point set in $U(Q)$ that has the most points, $Num(G)$ calculates the number of the points in point set G . “+” and “-” operations are equivalent to the corresponding operations in set theory.

A point set with proper interference is fitted by the least square method and the modified method respectively, the result is shown in Fig.11. The black points are the points to be fitted, the level part has 380 points, the slant part has 276 points, and the white line represents the fitted line. It can be clearly seen that the modified method can precisely find the optimal line and least square method can't.

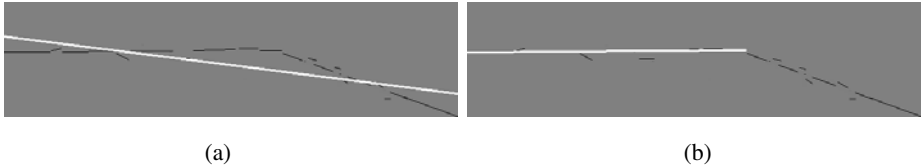


Fig. 11. Comparison of two methods: (a) The least square method; (b) The modified method

Before applying the modified method to fit the centerline in Fig.9, we can also make use of the information in the previous image, which is the X coordinate of the intersection of the slant parts in the previous image. According to this value, the Fig.9 can be divided into left and right parts. Then the modified method can be applied to the two parts respectively. The aim of separation is to improve the efficiency of this method as it is sensitive to the number and distribution of initial points. The result is shown in Fig.12. Four straight lines, a_1a_2 , a_3a_4 , a_5a_6 , a_7a_8 , are fitted and three main feature points can be calculated according to the four lines.

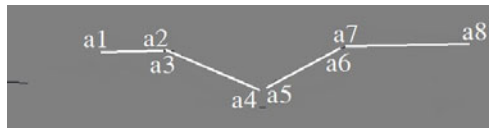


Fig. 12. Fitting line by the modified least square method

5 Conclusions

We develop a new image processing method for V-shape weld image and its effect has been verified by the weld image in this paper. Actually, median filtering, LOG filtering, threshold segmentation and particle filtering, aren't improved here, what we do is to arrange the order of these methods and set the parameters that are more suitable to our images. On the contrary, self-adapting dynamic window method and double local search method are proposed for improvements, comparisons with the methods used in literatures [6-9] have been described in section 2.3 and section 3. As the standard least square method can't find the optimal line from the centerline, we set the linearity as the principle to eliminate the noise points and refit iteratively so that the newly proposed modified least square method can have the ability of fitting the optimal line. The time consumed by applying the Standard Hough Transform used in literatures [8], [10-12] to process the centerline in this paper is almost five times as much as that consumed by applying the modified method on the same computer platform.

However, there are still some limitations in our image processing methods. In a few cases, the judgment used for double local search method is not right, consequently, the centerline may be incorrect and this can result in the wrong fitted lines and then wrong image processing result. The efficiency of the modified least square method is sensitive to the number of data points in the centerline, when it is applied to fit a relatively long centerline, i.e. having 2000 points, it may be time-consuming and has no advantages over Hough Transform. And the criterion for judging whether the image processing result is right or wrong should also be developed, as the result of the previous weld image decides the region of interest in the next one.

References

1. Sharif, L.H., Yamane, S., Hino, Y., et al.: Sensing and digital control of weld pool with visual welding robot [C]. In: The 26th Annual Conference of the IEEE Industrial Electronics Society, Nagoya, Japan, pp. 1521–1526 (2000)
2. Menno, D.G., Ronald, A., Ben, J., et al.: Real-time seam tracking for robotic laser welding using trajectory-based control. *Control Engineering Practice* 18(8), 944–953 (2010)
3. Shen, H.Y., Lin, T., Chen, S.B., et al.: Real-time seam tracking technology of welding robot with visual sensing. *J. Intell. Robot. Syst.* 59, 283–298 (2010)
4. Liu, Z.G., Chen, Z.X.: Research on image processing for laser sensing. *Electric Welding Machine* 39(4), 133–137 (2009)
5. Yu, D.H.: Image detection and processing technology. Xidian University Press, Xi'an (2006)
6. Li, Y., Xu, D., Shen, S., et al.: A image processing and features extraction method for structured light of welding seam. *Transactions of The China Welding Institution* 27(9), 25–30 (2006)
7. Li, K.: Research on vision-based tracking system for pipeline welding robot. Harbin Institute of Technology (2007)
8. Huang, W.Y.: Welding seam tracking system based on image sensor technology. Harbin University of Science and Technology (2007)
9. Chui, Y.B.: Research on the methods of welding seam images processing. Harbin University of Science and Technology (2006)
10. Opas, C., Guo, L.F.: A modified hough transform for line detection and its performance. *Pattern Recognition* 32, 181–192 (1999)
11. Yi, L.: Fast hough transform straight line detection based on classification. *Pattern Recognition* 23(11), 206–208 (2007)
12. Rozenn, D.: Statistical hough transform. *IEEE Transactions on Pattern Analysis And Machine Intelligence* 31(8), 1502–1509 (2009)

Melon Image Segmentation Based on Prior Shape LCV Model

Yubin Miao and Qiang Zhu

School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
ybmiao@sjtu.edu, cnzhuq713@126.com

Abstract. A new LCV (local chan-vess) model algorithm based on prior shape focusing on segmentation of melon image is proposed in this paper to measure the micro changes of morphological parameters. During the process, local boundary information image of melon is acquired firstly through mathematical morphology algorithm, and construct LCV model according to the form of traditional CV model. After that, a prior shape can be obtained by mathematical morphology and spline interpolation algorithms, and then be integrated into LCV model functional through a shape comparing function, thus the LCV model based on prior shape is constructed. Compared with traditional edge detection and segmentation algorithms, the new algorithm proposed in this paper could obtain more ideal boundary information.

Keywords: melon morphology, image segmentation, prior shape, LCV model.

1 Introduction

Machine vision technology has been widely used in plant growth test, quality inspection and non-destructive measurement, etc. And it's a key technology of how to get the precise target contour from the machine vision image. Common types of fruit images segmentation algorithm are: the color space to distinguish law, texture analysis, morphology analysis, wavelet analysis, multiple parameters combination analysis, images in different spectral bands identification method, and so on. Such as Zhang Yajing et al[1] has used Lab color model for the apple image segmentation. Zeng Qingbing et al[2] has proposed a watershed algorithm based on morphological processing for non-destructive measurement of fruit diameter of overlap grapes. Ying Yibin, Rao Xiuqin et al[3] have proposed fruit shape classification methods based on Zernike matrix. And there are other methods of fruit image segmentation which are based on wavelet transform, Hough transform algorithm and so on. The algorithms have a good effect when they are used to split the images that have great differences in color, uniform brightness, or obvious boundary. However, for the reason that the melon images have the characteristics of uneven brightness, and the melon edges overlapping with the leaves, stem and goity, the above methods won't get a satisfied result when they are used in the melon image segmentation.

In addition, in the field of image segmentation, the active contour models have become a very important model algorithm, widely used in image segmentation and

image understanding fields. Mumford et al[4] have proposed a segmentation model that combines image boundaries and regions. And based on that, Chan and Vese have proposed a simplified model which is the C-V model. The C-V model can effectively split the images that non-gradient defined. But the C-V model uses the global statistical information to split the images, and it won't split the local non-uniform images accurately. In recent years, it has been the two main methods for the researchers to improve the C-V model to blend the local information of the target image[6,7,8] and prior shape of the segmentation target[9,10] into pan function model. On this basis, the article proposes a LCV model algorithm based on prior shape, which is used to melon image segmentation and the exact extraction of the parameters of the morphological features in complex background.

2 Chan-Vese Model

C-V model is based on Mumford-Shah model, and it is a simplified image segmentation model of optimal regional division. It's main principle is to treat image as piecewise continuous gray-scale function. When the curve evolves to the edge of the detected target, the pan-energy function of the model will get a minimum, and the function as:

$$E^{CV}(c_1, c_2, C) = \mu \bullet length(C) + \nu \bullet Area(C) + \lambda_1 \int_{in(C)} |u_0(x, y) - c_1|^2 dx dy + \lambda_2 \int_{out(C)} |u_0(x, y) - c_2|^2 dx dy \tag{1}$$

In which, u_0 stands for the image to be split, C stands for any variable boundaries, c_1 and c_2 stand for gray-scale mean of the internal and external of boundary C respectively, μ , ν , λ_1 , λ_2 stand for weight factor. The first two are smooth entries to ensure the smooth of the contour C . The latter two are fitting items, so as to the evolution curve close to the ideal target profile.

In order to establish variational level set model, the article introduces the Heaviside function and Dirac function:

$$H(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases}, \delta(z) = \frac{d}{dz} H(z) \tag{2}$$

And the C-V energy model can be rewritten as:

$$E^{CV}(c_1, c_2, \phi) = \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy + \nu \int_{\Omega} H(\phi(x, y)) dx dy + \lambda_1 \int_{in(C)} |u_0(x, y) - c_1|^2 H(\phi(x, y)) dx dy + \lambda_2 \int_{out(C)} |u_0(x, y) - c_2|^2 (1 - H(\phi(x, y))) dx dy \tag{3}$$

The problem of seeking the minimum of the $E^{CV}(c_1, c_2, \phi)$ can be solved by solving the Euler-lagrange equation corresponding to pan-energy function. And finally we can get the level set evolution equation as follows:

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi) [\mu \operatorname{div}(\frac{\nabla \phi}{|\nabla \phi|}) - \nu - \lambda_1(u_0 - c_1)^2 + \lambda_2(u_0 - c_2)^2] \tag{4}$$

$$\phi(0, x, y) = \phi_0(x, y) \text{ in } \Omega \tag{5}$$

$$\frac{\delta_\epsilon(\phi)}{|\nabla \phi|} \frac{\partial(\phi)}{\partial \vec{n}} = 0 \text{ on } \partial \Omega \tag{6}$$

Gray-scale mean, c_1 and c_2 , can be updated at each iteration by the following way:

$$c_1(\phi) = \frac{\int_{\Omega} u_0(x, y) H_\epsilon(\phi(x, y)) dx dy}{\int_{\Omega} H_\epsilon(\phi(x, y)) dx dy} \tag{7}$$

$$c_2(\phi) = \frac{\int_{\Omega} u_0(x, y) (1 - H_\epsilon(\phi(x, y))) dx dy}{\int_{\Omega} (1 - H_\epsilon(\phi(x, y))) dx dy} \tag{8}$$

By the use of numerical difference, the ϕ^{n+1} after each iteration can be obtained as the approximation of the zero level set of the current moment.

3 LCV Model Based on Prior Shape

The traditional C-V model evolves based on the global statistical information, and it can not accurately separate local non-uniform images, so the LCV model based on a prior shape is proposed. This model mainly includes two steps: First, extract the local information of the image and put them into the C-V model, the formation of LCV model through the establishment of energy function, that's LCV model; The second is to extract the prior shape from the melon and to establish the signed distance function of the priori shape. The function is embedded into LCV model through the shape comparing function, that's LCV model based on prior shape.

3.1 Access to Local Image Information

Access to local image information mainly includes two steps: image pre-processing and morphological reconstruction.

The original image of melon contains background objects such as stem, leaves, pedicle and so on. They overlap and adhere to the edge of melon. So it's difficult to extract accurately the complete edge of the melon. Here to weaken the background through image sharpening, and try to highlight the impact of local melon details. In the

commonly used sharpening operator, the gradient operator can highlight the edge of melon and internal texture detail very well, so this paper firstly to process the melon RGB image to gray scale image, and then use the gradient method for image sharpening, as shown in Figure 1 (b).

To facilitate morphological image processing, the image G needs gradient threshold processing. That's setting a threshold based on the experimental experience. threshold, If $|G[f(x,y)]| \geq \text{thresh}$ $bw(x,y) = 255$; else $bw(x,y) = 0$; end. Where bw binary image obtained after thresholding. It's treated as an input in the morphological processing, as shown in Fig. 1(c).

And then use morphological reconstruction method to process the gradient image.

Mathematical morphology is based on a collection transformation method. This paper needs opening operation, closing operation, and morphological reconstruction to process the gradient image.

First, select appropriate structural elements for closing operation of the binary image bw . It can remove the details within melon and the operation results in large white areas of melon. Then remove the "black islands" that may appear inside the melon with the holes-filling method. Finally, remove the burr around the melon with open operation on the image, and get a rough area melon, as shown in Figure 1 (d).

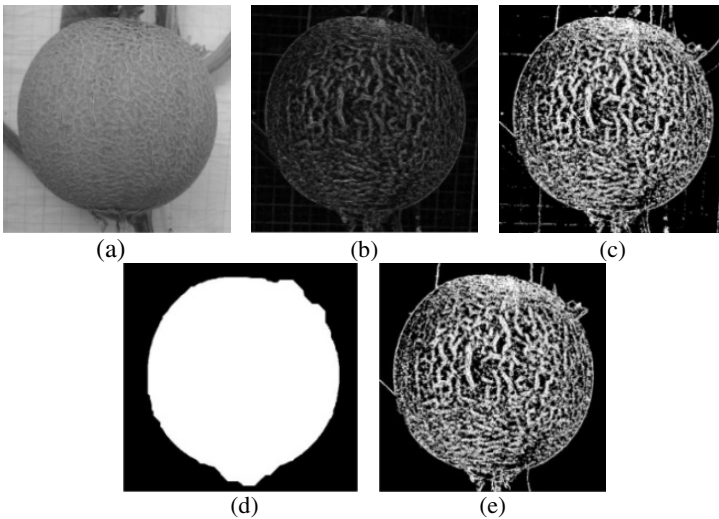


Fig. 1. Image preprocessing and morphology processing (a) gray-scale (b) gradient sharpening (c) threshold processing (d) opening and closing operations (e) reconstructed image

Finally, this paper use fig.1(d), treated as the mask image and fig.1(c), treated as the marker image for morphological image reconstruction. The result shows in fig.1 (e) that the reconstructed image gets a lot less background details than original image. Thus it highlights the local information around melon. So the algorithm in this study is based on this local information image.

The establishment of LCV model is still based on region-fitting method. Here is LCV energy model similar to the CV model:

$$\begin{aligned}
 E^{LCV}(d_1, d_2, \phi) = & \lambda_3 \int_{in(C)} |u_1(x, y) - d_1|^2 H(\phi(x, y)) dx dy \\
 & + \lambda_4 \int_{out(C)} |u_1(x, y) - d_2|^2 (1 - H(\phi(x, y))) dx dy
 \end{aligned}
 \tag{9}$$

Where u_1 is defined as the local information image, d_1 is the mean gray value within the area C , d_2 is the mean gray value without the area C . $\lambda_3, \lambda_4 > 0$ are the weighting coefficients. The numerical solution is basically the same with the solution of above CV model.

3.2 Construct Prior Shape Model

On the basis of the reconstructed image gotten before, extract rough boundary, remove false edge, and then evenly take points to spline interpolation, to get the prior shape of the melon. And then construct the signed distance function of the prior shape.

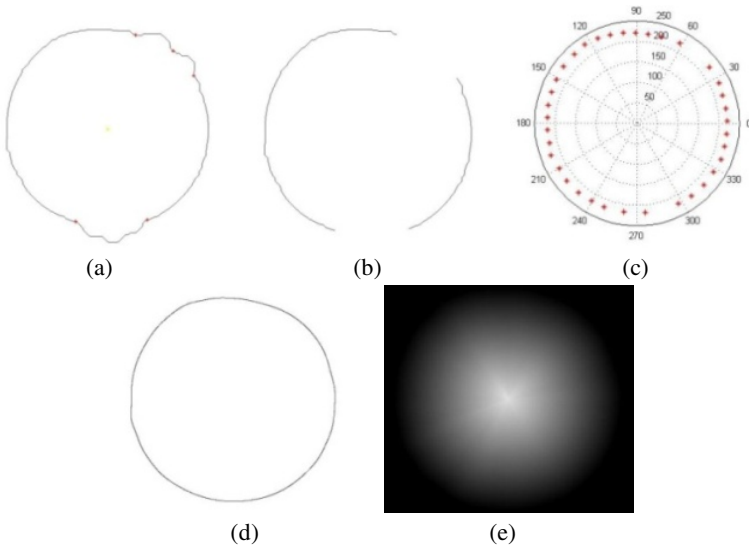


Fig. 2. Spline interpolation contour and prior shape: (a) rough contour (b) remove fake contour (c) uniform interpolation points (d) prior shape (e) signed distance image

3.2.1 Acquire Prior Shape

First, a rough contour can be obtained by making a open and a close operation to reconstructed image (fig. 2(a)). As shown in the figure, there are several bumps where stem and leaf overlap with melon edge, which are obviously fake edge information, need to be removed. Remove the fake edge by tracking extreme points first. Then choose interpolation points on the remained edge to do spline interpolation. A smooth edge curve of melon can be achieved as the prior shape of melon edge. As shown in fig.2 (d), the interpolation curve is much close to the real edge of melon. But there are

few wares generated because of too much or too little interpolation points, which makes part of the curve deviate from melon’s real edge.

3.2.2 Construct Signed Distance Function

Conventionally, the generation process of signed distance function is divided into to steps. Generate signed image first, then calculate the distance from points to curve. The closed curve divides the image into two parts. Define the interior as positive, exterior as negative, and on curve as zero. So the signed image is obtained. Then calculate the euclidean distance between every pixel to its nearest edge point. So the distance image is obtained. Finally, the signed distance function is achieved by multiply signed image by distance image.

Make full use of the data gotten from spline interpolation in the former step to simplify the building process of the signed distance function. Any point on the prior shape gotten from the interpolation has its own coordinates in polar coordinate $Q(\theta, r)$. First, calculate the polar coordinates $I(\theta, d)$ of each pixel with the melon core as the center. At the same time, the melon profile has a high roundness, and it can be approximated looked as a round. Therefore, it's apparently that the nearest boundary to the pixel point is $Q(\alpha, r)$. And that is the intersection of the boundary and the radius where $I(\alpha, d)$ stays. And so the Euclidean distance of the pixel point $I(\alpha, d)$, to its closest border point is $|r-d|$. With $f(x, y) = r-d$ as the signed distance function of the prior shape, it will well meet the characteristics of the signed distance function mentioned before. If $d > r$, it indicates the point $I(\alpha, d)$, is in the external of the curve, and $f(x, y) = r-d < 0$. And if $d \leq r$, it indicates the point, $I(\alpha, d)$, is in the internal of the curve or on the curve, and $f(x, y) = r-d \geq 0$. And this will be the initial signed distance map of the melon. As shown in figure 2 (e).

3.3 LCV Energy Model Based on Prior Shape

Traditional CV model, LCV model and prior shape is used in combination by this paper, and a new LCV model based on prior shape is constructed.

Aim to introduce the signed distance function of prior shape, a shape comparison function constructed by Cremers et.al.[11] is used. So, the evolution of energy model is restrained by shape comparison. The formula is:

$$E_{shape}(\phi) = \int_{\Omega} (\phi - \phi_{(0)})^2 dx \tag{10}$$

Obviously, the role of integral is to emphasize the prior shape in the whole range. In the formula, ϕ stand for image edge to be segmented, ϕ_0 stand for prior shape, both of them are signed distance function.

As a result, the new function of energy model is constructed as:

$$E(c_1, c_2, d_1, d_2, \phi) = E^{CV}(c_1, c_2, \phi) + E^{LCV}(d_1, d_2, \phi) + \beta E_{shape}(\phi) \tag{11}$$

After introducing Heaviside and Dirac function, the energy function is rewritten as:

$$E_{shape}^{LCV}(c_1, c_2, d_1, d_2, \phi) = \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy + \nu \int_{\Omega} H(\phi(x, y)) dx dy + \int_{in(C)} \lambda_1 |u_0(x, y) - c_1|^2 + \lambda_3 |u_1(x, y) - d_1|^2 H(\phi(x, y)) dx dy + \int_{out(C)} (\lambda_2 |u_0(x, y) - c_2|^2 + \lambda_4 |u_1(x, y) - d_2|^2) (1 - H(\phi(x, y))) dx dy + \int_{\Omega} (\phi - \phi_0)^2 dx \tag{12}$$

In the formula, ν and β stand for weight coefficient. Usually make $\nu = 0$, $\lambda_1 = \lambda_2$, $\lambda_3 = \lambda_4$. ϕ can be solved through solving the corresponding Euler-Lagrange equation of energy function. The solving process is same with formula (3)'s.

$$\frac{\partial \phi}{\partial t} = \delta_{\epsilon}(\phi) [\mu \operatorname{div}(\frac{\nabla \phi}{|\nabla \phi|}) - \nu - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 - \lambda_3 (u_1 - d_1)^2 + \lambda_4 (u_1 - d_2)^2 + 2\beta(H(\phi) - H(\phi_0))] \tag{13}$$

In the formula, ϕ_0 stands for signed distance function of prior shape, as constraint model of this algorithm. And using the core of the prior shape as center and the average radius of the prior shape as radius, draw a circle as the initialization curve of the evolution curve, which can restrain the curve to evolve near the prior shape. In the mean time, because the initialization curve is near to the real edge, less iteration times can make evolution curve achieve the real edge of melon.

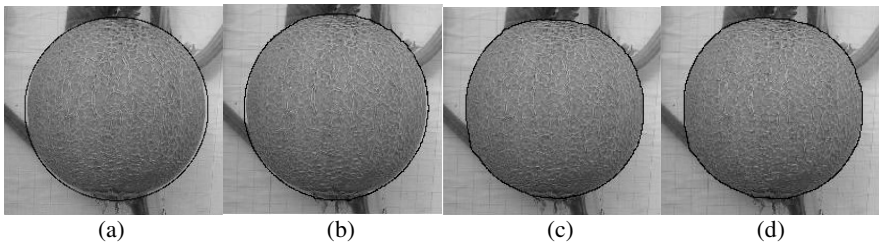


Fig. 3. Processing result with LCV model based on prior shape: (a) initialization (b) interpolation 20 times (c) interpolation 50 times (d) interpolation 100 times

Fig 3 shows the segmentation process of melon image by algorithm by this paper, he initialization curve is the circle which use the core of the prior shape as center and the average radius of the prior shape as radius. Curves in the figure stand for initialization curve and evolution curve iterated for 20 times, 50 times and 100 times.

4 Results and Analysis

Figure 4 shows the results of edges of melon detected by canny edge detection operator, morphological image processing, spline interpolation and CV model algorithm respectively. As shown in figure 4(a), detection result of canny contains a lot of interference information, from which ideal edge information can't be extracted. As shown in figure 4(b), morphological image processing can not remove stem and leave that overlaps with melon edge, which reduces the precision of edge detection of melon image greatly. As shown in figure 4(c), spline interpolation based on result of morphological image processing can remove interference information well, but, the curve show waves through interpolation, which makes the curve can not stick to the real edge of melon. As shown in figure 4(d) CV model divide stem and leaf area, which is similar with melon in gray, into melon area, leading to that stem and leaf can not be removed.

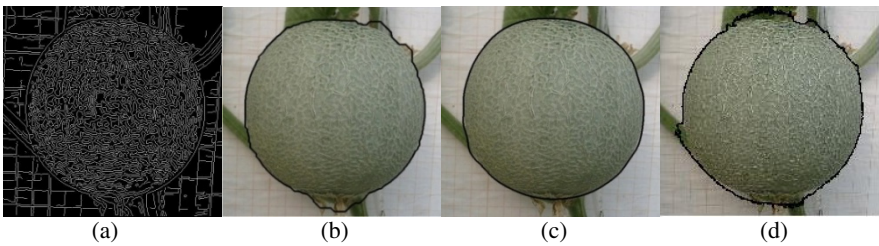


Fig. 4. The comparison of several image segmentation methods: (a) canny (b)morphological processing (c) Spline interpolation (d) CV model

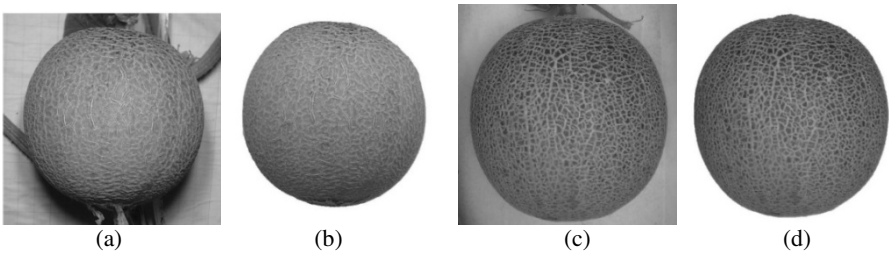


Fig. 5. Segmentation results of different melon image: (a)melon1 (b)segmentation of melon1 (c) melon2 (d) segmentation of melon1

Figure 5 shows the detection result of target melon edges by algorithm by this paper. As shown in the figure, algorithm by this paper is able to remove the fake edge information where stem and leaf overlaps with the melon edge, and obtain a accurate smooth contour curve of melon edge. Compared with traditional edge detection algorithm (canny e.g.) and CV model algorithm, algorithm by this paper can remove interference edge information well, and achieve a single smooth contour curve. Compared with morphological image processing and spline interpolation, algorithm by this paper can eliminate the waves of curve caused by interpolation through evolution,

which makes the final evolution curve keep to the real edge of melon well, and improve the accuracy of edge detection of melon greatly.

Table 1 shows the comparison between the measurement results of melon diameters by algorithm by this paper and vernier caliper in the same time.

Table 1. Evaluation of melon image processing algorithm accuracy

Figure No.	Fig3(a)	Fig5(a)	Fig5(c)
Measurement by caliper (mm)	105.12	109.75	119.47
Measurement by image (mm)	105.3846	110.0423	120.0848
Reference coefficient (pixels/mm)	18.3793	15.9934	14.6306
Relative error	0.25%	0.26%	0.51%

5 Conclusion

For the purpose of measuring the micro changes of morphological parameters of melon, a new LCV model algorithm based on prior shape is proposed. Through the image processing, acquire local boundary information image of melon through mathematical morphology algorithm first, and construct LCV model. Then, construct signed distance function of prior shape which is obtained by mathematical morphology and spline interpolation algorithms. Finally, integrate LCV model and signed distance function of prior shape to CV model, and construct a new LCV model based on prior shape. LCV model based on prior shape can keep the shape and slickness of curve through evolution, and decrease iteration times, shorten the evolution time, and the most important is improving the accuracy of acquiring melon edges.

Experiments show that, for melon images acquired in natural environment, algorithm in this paper can accurately segment and measure the longitudinal and horizontal diameters of fruit of melon, which can provide data reference for the analysis of the growth pattern of melon and fertilizer decision.

References

1. Zhang, Y., Li, M., Liu, G., Qiao, J.: Separating Adjoined Apples Based on Machine Vision and Information Fusion. *Transactions of the Chinese Society for Agricultural Machinery* 40(11), 180–183 (2009)
2. Zeng, Q., Liu, C., Miao, Y., et al.: Non-destructive measurement of diameter of overlapping grape fruit based on morphological image processing. *Transactions of the Chinese Society of Agricultural Engineering* 25(9), 356–360 (2009)
3. Ying, Y., Gui, J., Rao, X.: Fruit shape classification based on Zernike moments. *Journal of Jiangsu University (Natural Science Edition)* 28(1), 1–3 (2007)
4. Mumford, D., Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems. *Pure and Applied Mathematics* 42, 577–685 (1989)

5. Chan, F.T., Vese, A.L.: Active contours without edges. *IEEE Transactions on Image Processing* 10(2), 266–277 (2001)
6. Min, H., Wang, X.: A New Fitting Model Based on Boundary Information. *Electronics R&D* 11, 7–10 (2010)
7. Paragios, N., Deriche, R.: Geodesic active regions: A new framework to deal with frame partition problems in computer vision. *Int. J. Comput. Vis.* 46(3), 223–247 (2002)
8. Darolti, C., Mertins, A., Bodensteriner, C., et al.: Local region descriptors for active contours evolution. *IEEE Trans. Image Processing* 17(12), 2275–2288 (2008)
9. Wang, X., Zhao, Y.: Liver CT image segmentation based on prior shape CV model. *Journal of Optoelectronics Laser* 21(6), 953–956 (2010)
10. Duan, X., Xia, D.: Level Set Model for Segmentation of Cardiac MRI Image Based on Ellipse Shape Restriction. *Computer Engineering* 33(16), 14–17 (2007)
11. Cremers, D., Sochen, N., Schrorr, C.: Towards recognition based variational segmentation using shape priors and dynamic labeling. In: Griffin, L.D., Lillholm, M. (eds.) *Scale-Space 2003*. LNCS, vol. 2695, pp. 388–400. Springer, Heidelberg (2003)

An Improved Method for Terrain Mapping from Descent Images

Xiaoliang Xue¹, Cai Meng¹, and Yang Jia²

¹ School of Aerospace, Beihang University,
Beijing, 100191 China

² China Academy of Space Technology
Beijing, China

xue86buaa@gmail.com

Abstract. In the exploration of the planets of our solar system, images taken during a lander's descent to the surface of a planet provide a critical link for surface reconstruction. In use of descent images taken by landing camera is an effective way for terrain mapping. The method by dividing virtual planes for reconstruction has been improved in our paper. Based on plane homograph relationship, the match-related value noise is eliminated by gauss filter, and discontinuity disparity areas are smoothed by introduction of smooth-constraint match cost function. We demonstrate experimental results on synthetic images and our method can achieve better accuracy and effectiveness for terrain mapping.

Keywords: descent images, homograph, image match, terrain mapping.

1 Introduction

Future space missions that land on planet may include a downward-looking camera mounted on the vehicle as it descends to the surface. Images taken by such a camera during the descent provide a critical link between orbital images and lander images on the surface of the planet. With the help of inertial measurement unit and laser range, and by matching the descent images against orbital images, the descent vehicle can localize itself in global coordinates and therefore, achieve precision landing. Through analysis of descent images, we can build a multi-resolution terrain map for safe landing, rover planning, navigation, and localization. This paper addresses the issue of generating terrain map from a stereo descent images.

Terrain mapping from descent images is not yet widely discussed now. In literature[1], the conventional rectification process is replaced by a set of anti-aliasing image wrappings corresponding to a set of virtual parallel planes, and then correlates adjacent frames to match pixels for depth-recovery, however the match method is not precise enough by only implying SSD measure. [2]reconstruct depth values by recursive motion and orthogonal decomposition methods, but only base on sparse feature points, not the whole terrain mapping. Another method to get dense terrain is used in[3] by rectify the stereo image, yet the stereo is not based on descent images but on

wide baseline. When the viewed terrain is a planar surface, the motion-recovery problem is ill-posed, since translations parallel to the surface appear similar to rotations about axes parallel to the surface. The method above is not suitable for this issue.

Terrain mapping mainly has two parts: motion refinement and depth recovery. In literature[1], a set of virtual parallel planes is used to generate adjacent homograph frames and then correlates the frames to match pixels in SSD match measure. The method above is improved in our paper. After pixel match based on normalized cross-correlation method, the match-corresponding values in the same and different virtual planar are filtered by Gaussian filtering so that the noise can be eliminated from the sequence of match-corresponding values. And an additional constraint is added that supports smoothness by penalizing changes of neighboring disparities[4], the pixel-wise cost and smoothness constraints are expressed by defining the new cost function, so we can get more smooth disparity map.

In the next three sections, we describe our method in details. The first section is the motion refinement by extracting SIFT feature points; then the depth-recovery process is discussed in the second section. At last we discuss our experiments on synthetic descent images.

2 Motion Refinement

It is unrealistic to expect the onboard inertial sensors to track the camera orientation with high precision, so motion refinement should be recovered from two or more frames. The matched features provide a rich set of observations to constrain the camera motion. In our method, we extract SIFT[5] features that is invariant to image scale change, then exclude external point by Ransac[6]. We get precise match points finally by satisfying pole constraint relationship.

The objective of motion refinement is to establish the precise camera motion between two adjacent frames such that the epipolar constraints are satisfied to subpixel accuracy. In [1] the author take the distance of predicted and observed feature locations as objective function. In addition, the depth value for every feature points is also included as optimization parameters, we find this method can not convergence good enough. We improve the optimization using the geometric constraints relationship[7]. d_i and d_i' is the feature point distance from the corresponding pole line.

$$\text{Min} \sum_{i=1}^n (d_i^2 + d_i'^2) \quad (1)$$

In our method, we take the camera coordinate system corresponding to the upper image as the world coordinate system. Let us assume that the camera motion between the two images is composed of a translation T and rotation R (3*3 matrix). We perform nonlinear minimization using the Levenberg-Marquardt algorithm[8] with the estimated position as the starting point.

3 Depth Map Recovery

3.1 Depth Recovery Principle Using Homography

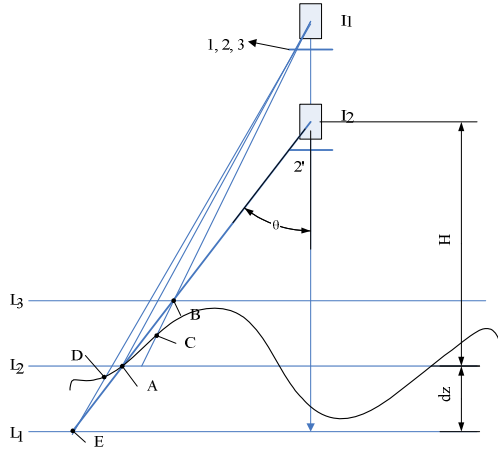


Fig. 1. Depth recovery diagram using homography

Fig1 shows the depth-recovery principle using a set of virtual planar surfaces. If we designate the terrain reference point A on surface L2 and the corresponding pixel 2' in the second image I2. The points labeled 1,2,3 in the first image I1 are generated by homography of virtual surfaces L1,L2,L3. The real match pair is (2, 2'), we can use pixel match restraint to eliminate the mismatch pair(1, 2') and (3, 2'). So we can determine the corresponding surface to pixel 2' is L2, therefore the point depth is determined.

We can determine the homography using the camera internal parameters K and external parameters(R,T)[9]. The distance from coordinate origin to the virtual plane can be replaced by the depth value of the plane. If we assume that the reference image defines I1, the registration image defines I2, the divided depth levels defines m, the maximize and minimize depth value as z_max and z_min. The depth value of every virtual plane is that:

$$z_k = z_{min} + k * (z_{max} - z_{min}) / m (k = 0 \dots L-1) \tag{2}$$

where H_k is the homograph of the surface whose depth value is z_k

$$H_k = K(R + t * n_d^T)K^{-1}(n_d = n / z_k) \tag{3}$$

Where the unit normal vector of the virtual plane is n = (0,0,1)^T and the homograph point is x_2(k)

$$x_2(k) = H_k x_1 (k = 0 \dots L-1) \tag{4}$$

We define D as pixel-match measure value, the virtual plane h whose pixel-match measure value is the smallest in levels L is the point depth value

$$D_h = \text{Min}D_k(x_1, x_2(k)) \quad (5)$$

After determine the point depth, we can easily get the three-dimensional coordinates of the point.

3.2 Improved Method for Matching Strategy

3.2.1 Gaussian Filter for Match Similarity Value

In ideal situation, the estimated depth value at each pixel is the depth of the plane whose corresponding NCC pixel match $D_k(x_1, x_2(k))$ is the smallest. However, as the uncertain influence of the pixel match method and the wide disparity search area, some noise can be included in the pixel match sequence. As for that we use Gaussian filter preprocessing to reduce the noise.

Based on the depth continuity in neighboring area, the NCC pixel match score not only in the same virtual surface but in different virtual surfaces should be continue. The origin pixel match sequence D corresponding to different virtual surfaces can be filtered into sequence D_1 . We use the 1D Gaussian filter $g_1(w_1, \sigma_1)$, where (w_1, σ_1) is the window size and variance of the filter.

$$D_1 = D \otimes g_1(w_1, \sigma_1) \quad (6)$$

After that, in the same virtual surface we apply the 2D Gaussian filter $g_2(w_2, \sigma_2)$ in each pixel neighboring area, where (w_2, σ_2) is the window size and variance of the 2D filter.

$$D_2 = D_1 \otimes g_2(w_2, \sigma_2) \quad (7)$$

3.2.2 Aggregation of Costs

Pixelwise cost calculation is generally ambiguous and wrong matches can easily have a lower cost than correct ones, due to noise, etc. As discussed in[4], an additional constraint is added that supports smoothness by penalizing changes of neighboring disparities. So the pixelwise cost and the smoothness constraints are expressed by defining the function $L_r(p, d)$ that depends on the neighboring disparity at each pixel

$$\begin{aligned} L_r(p, d) = & D(p, d) + \min(D_r(p-r, d), \\ & D_r(p-r, d-1) + P_1, D_r(p-r, d+1) + P_1, \\ & \min_k D_r(p-r, k) + P_2) - \min_k D_r(p-r, k) (k = 0 \dots L-1) \end{aligned} \quad (8)$$

The origin matching costs from all directions that we use 4*4 neighboring area contributes for the new matching cost $L_r(p, d)$. L refers to one of the directions in the area, $D(p, d)$ is the match cost at pixel location p in the virtual surface d . We add a constant penalty P_1 for all pixels whose corresponding virtual surface changes a little

bit(i.e. 1) and a larger constant penalty P_2 , for all larger disparity. $\min_k D_r(p-r, k)$ is the smallest pixel match cost at pixel location (p-r) in all L virtual surfaces. The cost $L_r(p, d)$ are summed over paths in all directions r.

$$S(p, d) = \sum_r L_r(p, d) \tag{9}$$

The smallest $S(p, d)$ in L virtual surfaces correspond to the correct virtual surface whose depth value is also the pixel depth. To further refine the depth value, the underlying cost curve can be interpolated by a quadratic curve[10] and the ‘subpixel’ depth value can be computed.

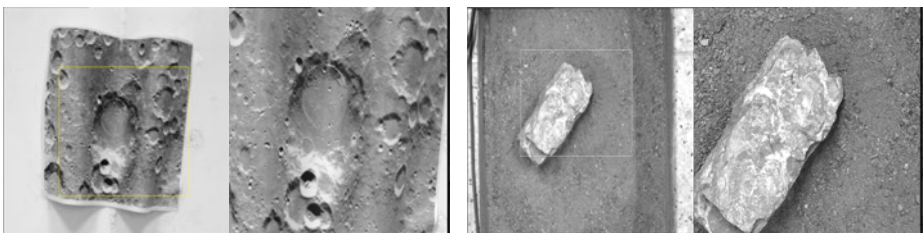
4 Experiments

In our experiment we use synthetic images including bunker stereo images and moon-texture stereo image. The image size is 768*576. We construct the common image area that can be determined by affine warping, that is the box marked area in figure2. Bunker stereo images are gathered by controlling the camera down to the surface. Here are our experiment analysis and result below.

4.1 Compare with Different Match Methods

We find that different match methods can have depth images with big difference. So we compare the depth images generated by three match method: local cross-entropy[11], NCC method, SSD method. Figure 3 shows the each depth image result in our method with the match window 5*5, the images are generated before median filter.

As showed in fig3, the depth image corresponding to NCC method is the best, the SSD method result shows the worst and local cross-entropy depth image has mass depth discontinues holes.



a moon-texture stereo images

b bunker stereo images

Fig. 2. Stereo images for experiment

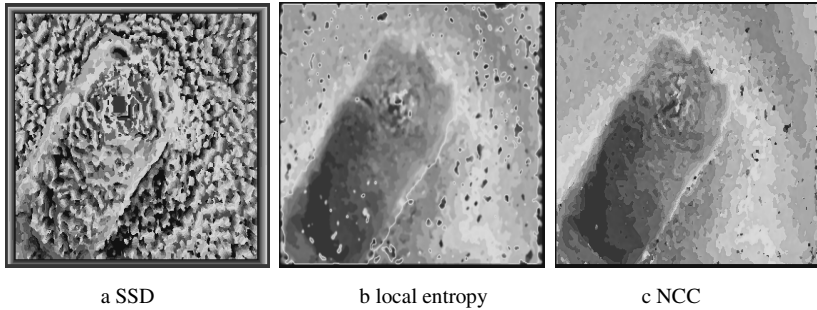


Fig. 3. Depth image contrast in different match methods

4.2 Recovery Result Contrast

In our experiment our method and the method in[1] are tested on the images showed in fig6 by comparing the 3D texture mapping. We use NCC match method with window size 5*5 and other parameters are the same. As showed in fig4, the left and right side separately correspond to the result in our method and the method in[1]. We can see clearly mass depth hole and rag are appeared in fig4a, that will make confused for real terrain information. Fig4b is our method result, compare with the fig6a, the texture mapping becomes more smooth and accurate.

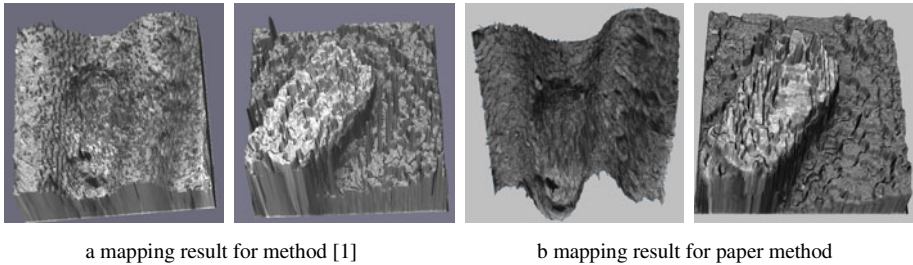


Fig. 4. Depth recovery result contrast in different methods

Also from fig4b, we can see the recovery result is fit with the real environment, for example, the tuber in the stone area in fig2b and arch area in the moon-texture area in fig2a.

4.3 Accuracy Contrast

In this part, we need test our method accuracy, including the match precision and the 3D location precision. For the match part, we take the SIFT points as the match reference points as correct match location, the match location error in our method and the method in[1] are calculated base on the SIFT location. The statistics error curve is showed in fig5:

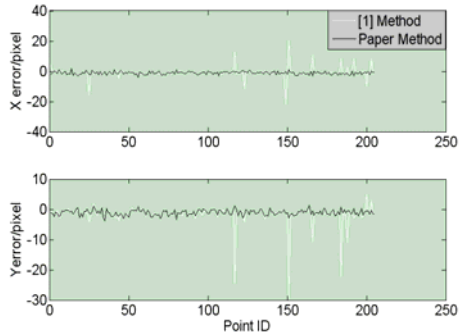


Fig. 5. Depth recovery result contrast in different methods

From fig5, we can see our method curve removes the burr generated by method[1], because of big mismatch error. The error mean and variance contrast are showed in talbe1 below:

Table 1. The match error contrast between two methods

Error(pixel)	Literature[1] Method	Paper Method
X direction mean	-0.9077	-1.1163
X direction variance	10.1779	0.9130
Y direction mean	-1.6569	-1.0906
Y direction variance	10.5577	0.8323

As showed in talbe1, Our match method has less variance due to the introduction of smooth constraint. So the match algorithm accuracy has been improved.

For checking the recovery accuracy, we select chess board as our experiment images, which are taken at the distance of 120cm and 80cm toward the .axis direction. The chess image size is 4*6. We compute the 3D coordinate of the chess-corners whose image locations are extracted before. We take the distance between the 3D corners and the fitting plane from the 3D corners as the recovery error. The error figure is showed in fig6.

In ideal situation, the distance between the corner to the fitting plane is zero. The statistics shows that in our method the mean and the variance of the distance error is [1.2mm, 1.5mm]., and the error in [1] is [2.5mm,4.5mm]. The distance error range is 4mm, and the error range in[1] is between -10mm to 8mm.

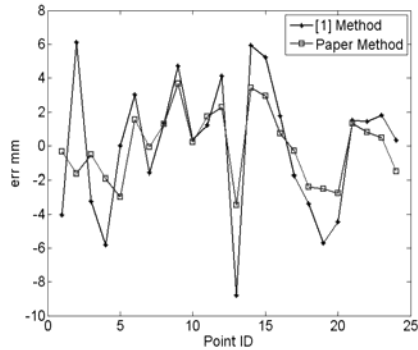


Fig. 6. Depth recovery result contrast in different methods

5 Summary

We have presented an improved method for extracting depth maps from descent images. Our improved work mainly include the two points below:

- 1) By rectifying the resolution of the two images, the image match process is only in the common area, not the whole image, that decreases the computer time.
- 2) After Gaussian filter preprocessing to the pixel match cost, an additional constraint is added that supports smoothness, so that the disparity image can be more smooth.

The experiments have resulted in maps that our method can get smooth disparity and more accurate reconstruction score than method in[1]. Our method can be effectively used for navigation and planning in terrain mapping.

References

1. Xiong, Y., Olson, C.F.: Computing depth maps from descent images. *Machine Vision and Application* 16, 139–147 (2005)
2. Shao, W., Sui, S.: Two algorithms and error analysis for 3D terrain reconstruction of small body. *2010 International Conference on Artificial Intelligence and Computational Intelligence* 1, 492–496 (2010)
3. Ming, J., Xianlin, H.: Lunar terrain mapping based on stereo vision. *Journal of natural science of HeiLongjiang University* 24(5) (2007)
4. Hirschmuller, H.: Accurate and efficient stereo processing by semi-global matching and mutual information. In: *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 807–814 (June 2005)
5. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2), 91–110 (2004)
6. Fischler, M.A., Bolles, R.C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communication of the ACM* 24, 381–395 (1981)

7. Zhang, Z., Deriche, R., Faugeras, O.: A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence Journal* 78, 87–119 (1995)
8. Levenberg, K.: A Method for the Solution of Certain Problems in Least Squares. *Quarterly of Applied Mathematics* 2, 164–168 (1944)
9. Fu, C., Wu: *Mathematical methods in computer vision*, pp. 68–71. Scientific Press (2008)
10. Xiong, Y., Matthies, L.H.: Error analysis for a real-time stereo system. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1098–1093 (1997)
11. Jiang, H., Shen, Z.: Matching tracking algorithm based on local cross-entropy. *Infrared and Laser Engineering* 34(6), 729–732 (2005)

Best View Selection of 3D Object Based on Sample Learning

Zhi Liu¹, Yipan Feng¹, Qihua Chen², and Xiang Pan¹

¹ College of Computer Science and Technology, Zhejiang University of Technology,

² College of Mechanical Engineering, Zhejiang University of Technology,
310023 Hangzhou, China
lzhi@zjut.edu.cn

Abstract. In connection with user perception-related features in the best view of 3D object, a method of best view selection of 3D object based on sample learning is proposed. Firstly, the AdaBoost algorithm is applied on the sample set for supervised learning in order to get the best view space of 3D object. Then, the rendering views come from different viewpoints in 3D object are compared and analyzed with the best view in shape similarity matching. The rendering view with highest similarity will be regarded as the best view of 3D object. The experiments show that the best view selected is very consistent with the user's subjective choices. This method has good validity and reliability.

Keywords: 3D object, best view selection, sample learning, similarity matching.

1 Introduction

The real world is composed of 3D objects, but with the limitation of imaging principle in human visual system, 3D objects only can be cognized after they are mapped into 2D views. The 2D views can maintain the features of 3D models depending on the location of viewpoints. So how to choose appropriate observation viewpoint for 3D models and how to measure the goodness of a view become difficult problems in the research of 3D model recognition.

In order to get important view characteristics, a series of evaluation criteria has been proposed to quantify the visual habits in observation objects of human beings[1]. This standard is divided into two categories: one is to analyze 3D models and form the best evaluation criteria[2] [3] [4]. Another method is to analyze and compare characteristics of the projection view and get best viewpoint. The significant advantage in this kind of method is all feature analysis is in 2D discrete space[5]. Nowadays, there are many methods of viewpoint selection and it has made great progress in this field. But there is still no suitable way for all kinds of objects. Depending on actual applications, different methods are selected to achieve good results.

Here, a method of best view selection of 3D object based on sample learning is proposed. In this way, user should define the best views according to user knowledge in the 3D model database for some kind of objects firstly. These best views will be used as the best view samples. Then the AdaBoost algorithm is applied on training

sample set for supervised learning in order to get the best view with user experience. In this model, the problem of best view selection is converted to the problem of geometric similarity in the shape space. The algorithm is showed as below Fig.1.

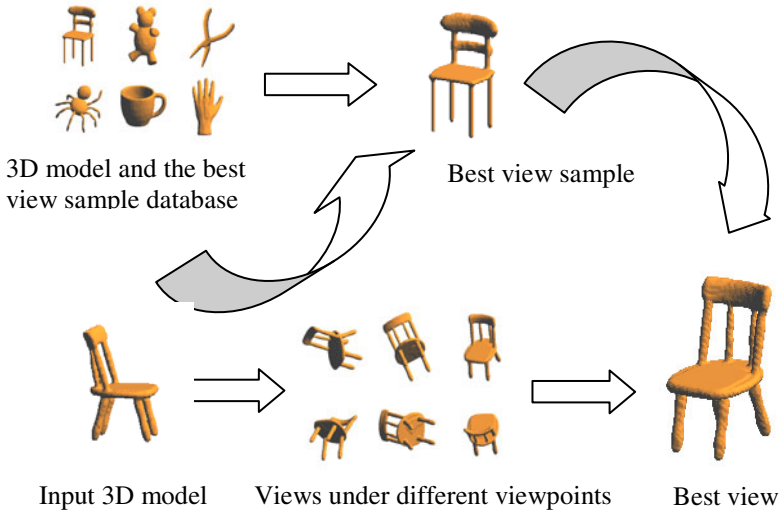


Fig. 1. The process of best view selection

2 Extraction of Best View Sample

According to the algorithm, the best view sample will be extracted based on the shape feature of 3D model firstly. The main steps include extracting shape feature, classifying through AdaBoost algorithm, and get the best view sample.

In order to make the feature value with geometric invariance on translation, rotation, and scaling, the 3D model will be preprocessed before feature is extracted. The PCA method is used to make normalization operation. So the multi-feature fusion method [6] is used to describe the features of 3D model. According to the strength of feature description ability, depth buffer feature with 186 dimensions, silhouettes feature with 150 dimensions, and ray based feature with 136 dimensions are selected to form feature descriptor DSR472.

Here, the AdaBoost algorithm is used on sample sets for supervised learning in order to establish the calculation model to obtain the best view sample through 3D shape feature. In our work, more than two types of model need to be categorized. So multi-class classifier need to be constructed. In general, the multi-class problem is converted to multiple binary-class problems. Literature[7] proposed two kind of strategies to extend a binary-class classification problem to a multi-class classification problem: one-against-one classification strategy and one-against-all classification strategy. In this paper, the one-against-all classification strategy is applied. So the multi-class classification can be obtained.

For an input 3D model, when the feature of the 3D model is inputted, the category label belonging to the 3d model can be gotten. In the best view sample database which is constructed according to user's criterion, the number of best view sample is equal to the number of categories of 3D model and each model has only one best view sample corresponding. So we can get the label of category from the shape feature of 3D model and then get the best view sample based on the category label.

3 Best View Selection

After obtaining the best view sample, the issue of selection of best view for the input 3D model is converted to geometric similarity issue in shapes space. For the input 3D model, the different rendered views coming from different viewpoints are compared with the best view sample on the shape similarity analysis. The geometry similarity is taken as a measure for the merits standard for the viewpoints. The view with highest similarity will be the best view of the model.

3.1 Description of View Feature

In this paper, a limited set of the boundary points is used to describe the target shape. The inner distances and inner angles are used to describe the features of 2D view[8] and the shape histogram of the boundary points is used to represent the feature of the target shape.

Firstly, the 2D view is processed to be a binary view by binarization processing. Then the Canny edge detection operator is used to extract the outline of binary view; The sampling rule[9] is used to sample outline and get the set of boundary points.

In this paper, the definition of the context of the boundary point shape[10] is extended. The distance parameter and direction parameter between the boundary points are changed. The Euclidean distance between two points is replaced as inner distance and relative direction between two points is replaced as inner angle. The log-polar transformation is used to describe the relationship between various border points and the whole contour. For the set of sample points (p_1, \dots, p_n) , each point is calculated with the remaining $n - 1$ points to form the shape histogram according to the formula (1). So each point corresponds to a feature vector and the whole target shape can be represented with the feature matrix.

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\} \quad (1)$$

In which, $k = 1, \dots, K$, K represents the product of direction parameter and distance parameter. The $(q - p_i) \in \text{bin}(k)$ represents that point q is belong to the k -th component of shape histogram relative to the point p_i .

3.2 Measure the Pros and Cons of Views

Here, the χ^2 [10] is used to calculate the matching cost between point p_i and point q_j . The cost function is showed as formula (2).

$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{1 \leq k \leq K} \frac{[h_{A,i}(k) - h_{B,j}(k)]^2}{h_{A,i}(k) + h_{B,j}(k)} \quad (2)$$

In which, $h_{A,i}(k)$ represents the shape histogram of point p_i of target A. The $h_{B,j}(k)$ represents the shape histogram of point q_j of target B. According to formula (3), the cost matrix between two target shapes is obtained and is recorded as C with size of $n \times n$.

Then, based on the cost matrix C , the matching operation of points can be processed and make the equation (3) to obtain the minimum value.

$$H(\pi) = \sum_{1 \leq i \leq n} C(p_i, q_{\pi(i)}) \quad (3)$$

Based on the analysis above, the points matching problem is attributed to the a typical two-way graph matching problem[11]. It can be solved through dynamic programming algorithm. Now, the similarity of two targets is represented with a non-vector. Because it is calculated based on the cost matrix, so that the greater the result, the less similar the targets. The smaller the result, the more similar the targets.

4 Experimental Analysis

In order to verify the effectiveness of the method, the experimental result is analyzed. The MSB 3D shape benchmark database supported by McGill University is used as the test database[11]. There are 19 model classes and 457 3D models in this database. Each type of model has 20-30 models. In this work, user should define 19 best view samples for 19 types of models firstly. That means one type of models will correspond to one best view sample.

Then the models in the database are divided into the training set and testing set. Here, 229 models in the database are selected to make supervised learning with AdaBoost algorithm. The other 228 models are used to test the algorithm.

The matching degree of visual perception is mainly based on whether the calculated result is met the user's visual perception or not. Here, the geometric similarity between orientation view and best view is represented as matching cost. It is showed in Fig.2. In this paper, the viewpoint with minimal matching cost is taken as the best viewpoint and the viewpoint with maximal matching cost is taken as the worst viewpoint. In Fig.2, several views on the best viewpoint and worst viewpoint are given and geometric similarity of these views with the best view sample is given too. The models on the top row come from the best viewpoints and the models on the following row come from the bad viewpoints.





















 64.9885	 46.1592	 62.5421	 56.5635	 43.8290
 75.1386	 61.6101	 92.0092	 77.1422	 71.9796
 74.8689	 61.0642	 56.7570	 85.3477	 84.4820
 94.2596	 84.5368	 70.1177	 96.8576	 92.4679

Fig. 2. The best viewpoints/bad viewpoints with the matching cost

We also select some 3D models from the PSB [12], such as turtle, horse, bucket and other models. These models are beyond the scope of the training set, but the best views of these models found in this method still have good visual perception.

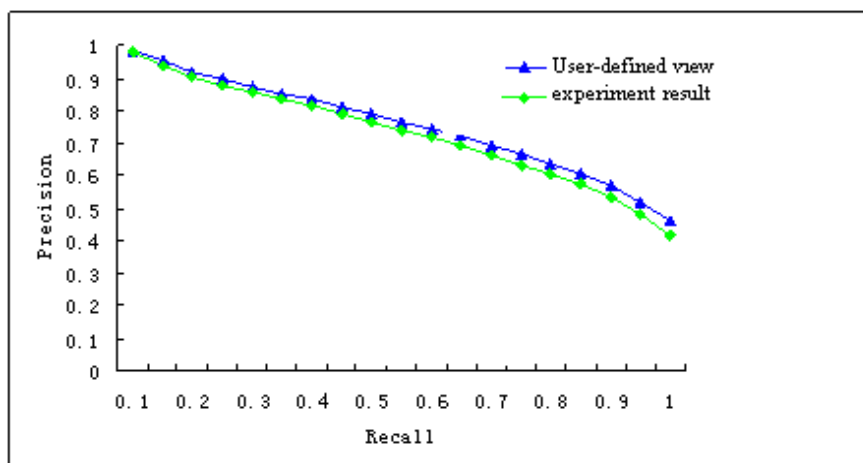


Fig. 3. The comparison between experiment result and user-defined result

In order to illustrate the effectiveness of this method proposed in this paper, the testing data will be evaluated through quantitative analysis. Precision-Recall curve is applied as quantitative appraisal index. The PR curve is showed in Fig.3. The two retrieval rate curves are so closed. That means the best view gotten from this method in this paper is very similar with the user-defined best view. It further proves the validity and accuracy of this algorithm.

5 Conclusions

Nowadays, there are still no unified evaluation criteria for the best view. The experiments show that the selection of best view selection of 3D object based on sample learning proposed in this paper is rather effective. The retrieval accuracy is used for statistical analysis and it can effectively describe the applicability of the algorithm. In next research work, the evaluation criteria for the best view need to be studied and defined in order to form a benchmark testing method. It will be used to analyze and evaluate different selection algorithms of best view.

References

1. Polonsky, O., Patance, G., Biasotti, S.: What's in an image: Towards the Computation of the Best View of an Object. *The Visual Computer* 21(8), 840–847 (2005)
2. Feixas, M., del Acebo, E., Bekaert, P., et al.: An information theory framework for the analysis of scene complexity[J]. *Computer Graphics Forum* 18(3), 95–106 (1999)
3. Yang, L., Wang, W., Wu, E.: Viewpoint Selection by Feature Measurement on the Viewing Plane. *Journal of Computer-aided Design & Computer Graphics* 20(9), 1097–1103 (2008)
4. Cao, W., Hu, P., Li, H.: Canonical Viewpoint Selection Based on Distance-Histogram. *Journal of Computer-aided Design & Computer Graphics* 22(9), 1515–1521 (2010)
5. Vázquez, P.P.: Automatic view selection through depth-based view stability analysis. *The Visual Computer* (25), 441–449 (2009)
6. Bustos, B., et al.: Automation Selection and Combination of Descriptors for Effective 3D Similarity Search. In: *IEEE Sixth International Symposium on Multimedia Software Engineering*, pp. 514–521 (2004)
7. Hao, W., Luo, J.: Generalized multiclass adaboost and its application to multimedia classification. In: *IEEE Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop* (2006)
8. Ling, H., Jacobs, D.W.: Shape Classification Using the Inner-Distance. *IEEE Transactions on Pattern Analysis and Matching Intelligence* 29(2), 286–299 (2007)
9. Tu, C., Diao, L., Lu, M., et al.: The Typical Algorithm of AdaBoost Series in Boosting Family. *Computer Science* 30(3), 140–145 (2003)
10. Belongie, S., Malik, J., Puzicha, J.: Shape Matching and Object Recognition Using Shape Context. *IEEE Transactions on Pattern Analysis and Matching Intelligence* 24(24), 509–522 (2002)
11. Scoot, C., Nowak, R.: Robust contour matching via the order-preserving assignment problem. *IEEE Transactions on Image Processing* 15(7), 1831–1838 (2006)
12. Agathos, D.: McGill 3D shape benchmark shape analysis group[EB/OL] (December 14, 2005), <http://www.cim.mcgill.ca/~shape/benchMark>

An Improved Retina Modeling for Varying Lighting Face Recognition

Yong Cheng¹, YingKun Hou², and Zuoyong Li³

¹ School of Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China

² School of Information Science & Technology, Taishan University, Taian, Shandong 271021, China

³ Department of Computer Science, Minjiang University, Fuzhou 350108, China

Abstract. An improved retina modeling based on bilateral filter is developed for face recognition under variable lighting. Bilateral filter is applied to estimate the adaptation factor X_0 (local lighting) in the Naka-Rushton equation, which models nonlinear processing of photoreceptors and outer plexiform layers. Difference of Gaussians filter (DoG) is also used to enhance image contours so as to model inner plexiform layer. Experimental results on the Yale B and CMU PIE face databases indicate the effectiveness of the proposed method.

Keywords: Face recognition, retina modeling, bilateral filter, varying lighting.

1 Introduction

Over the last many years, face recognition has made significant progress. The performance of automatic face recognition systems under controlled environments is usually satisfactory, however, there are still many challenges in uncontrolled environments. Small variations of lighting conditions often cause significant changes on face appearance, which could seriously degrade the face recognition performance. Although many methods have been proposed to solve this problem, such as Histogram Equalization (HE), Gamma Correction (GC), Multiscale Retinex (MSR) [1], Self Quotient Image (SQI) [2] and so on, varying lighting effects cannot be completely removed by these methods.

Recently, some methods inspired by human vision system (HVS) have been developed for image processing and feature extraction. Meylan et al. [3] presented a tone mapping algorithm, which was derived from a model of retinal processing. Inspired by Naka-Rushton equation they modeled the retinal photoreceptors' nonlinearity as two consecutive nonlinear operations and showed a good improvement in contrast while avoiding halos and maintaining good global appearance. Later, Vu et al. [4] proposed a retina modeling based lighting normalization method by using the two consecutive nonlinear operations and Difference of Gaussians (DoG) filter which also showed good performance on varying lighting face recognition. In the Naka-Rushton equation [3, 5] $Y=X/(X+X_0)$, X represents the input light intensity, X_0 is the adaptation factor, and Y is the adapted signal. In the original formulation [5], the X_0 is determined by the average light of entire reference field (an image) and it is a fixed

value for each pixel. Meylan et al. [3] adjusted the adaptation factor X_0 as a local variable given by the average light intensity in the neighborhood of one pixel, which is computed for each pixel by performing Gaussian filtering on its neighborhood. However, this imitating Amacrine cells and Horizontal cells by using Gaussian filtering for local lighting often produce inaccurate local lighting on the edges of an image. This is because Gaussian filtering carries out a simple weighted average of a neighborhood without considering intensity difference in this neighborhood. Inaccurate local lighting will cause distorted results in the later process of lighting normalization.

Bilateral filtering replaces a pixel's value by weighting the neighboring pixels considering both their geometric closeness and photometric similarity. It can preserve image details while getting image low-frequency information and get more accurate local lighting than Gaussian filtering does. Hence, in the improved retina modeling for variable lighting face recognition, we apply bilateral filtering to obtain the key parameter X_0 instead of Gaussian filtering. Experimental results on the Yale B and CMU PIE show that our method is highly insensitive to lighting variation.

2 Retina Modeling for Varying Lighting Face Recognition

2.1 Retina Model

The primate's retinal model are three functional layers, namely, the photoreceptors layer, the Outer Plexiform Layer (OPL) and the Inner Plexiform Layer (IPL) [6, 7].

The photoreceptors layer is composed of two cells: cones and rods. It is responsible for visual data acquisition and is related with a local image illumination compression respecting the illumination of neighborhood, which is achieved by the next cellular network: the Horizontal cells. This local compression has been modeled by the authors in based on the Naka-Rushton equation, and the model has been used for illumination normalization.

The Outer Plexiform Layer includes Bipolar and Horizontal cells. The functions of the Bipolar cells are simple to transmit signal from the OPL to IPL. The Horizontal cells are responsible for obtaining the illumination of neighborhood (local illumination). It feed back photoreceptors layer for local compression of image illumination. In [3] and [4], the local illumination was achieved by Gaussian filtering.

The Inner Plexiform Layer consists of the Ganglion cells and the Amacrine cells. This layer has two functions. The first is to modify illumination like the photoreceptors layer does. In this stage, the Amacrine cells provide local illumination like the Horizontal cells do. The second is to perform contour enhancement. This can be modeled by a spatial band-pass filter.

2.2 Bilateral Filtering

In 1998, Tomasi and Manduchi [8] proposed the bilateral filtering, which can smooth images while preserving edges by means of a nonlinear combination of some pixel values. It replaces a pixel's value by weighting the neighboring pixels considering their geometric closeness and photometric similarity, however, Gaussian filtering simply weights the neighboring pixels. A simple and important bilateral filtering is shift-invariant Gaussian filtering, in which both geometric closeness and photometric

similarity are measured with Gaussian functions of the Euclidean distance between their arguments. The bilateral filtering can be defined as following:

$$B(\vec{\chi}) = \delta(\vec{\chi})^{-1} \int I(\vec{\gamma}) \exp\left\{-\frac{|\vec{\gamma} - \vec{\chi}|}{2\sigma_d^2}\right\} \exp\left\{-\frac{|I(\vec{\gamma}) - I(\vec{\chi})|}{2\sigma_s^2}\right\} d\vec{\gamma}. \tag{1}$$

$$\delta(\vec{\chi}) = \iint \exp\left\{-\frac{|\vec{\gamma} - \vec{\chi}|}{2\sigma_d^2}\right\} \exp\left\{-\frac{|I(\vec{\gamma}) - I(\vec{\chi})|}{2\sigma_s^2}\right\} d\vec{\gamma} \tag{2}$$

where $I(\vec{\gamma})$ is the input image, $B(\vec{\chi})$ is the filtered image, $\exp\left\{-\frac{|\vec{\gamma} - \vec{\chi}|}{2\sigma_d^2}\right\}$ and $\exp\left\{-\frac{|I(\vec{\gamma}) - I(\vec{\chi})|}{2\sigma_s^2}\right\}$ is two Gaussian functions which are used for describing geometric closeness and photometric similarity, respectively. σ_d^2 and σ_s^2 are variances of the two Gaussian functions. $\delta(\vec{\chi})$ is the normalization factor.

2.3 Improved Retina Modeling

The illumination normalization algorithm based on retina model can be described as follows,

Step 1. The first local lighting compression imitates the Horizontal cells: The original image M whose intensities are normalized to $[0, 1]$, is processed by the following equations,

$$P(x, y) = (M_{\max} + A_M(x, y)) \frac{M(x, y)}{M(x, y) + A_M(x, y)}. \tag{3}$$

$$A_M(x, y) = B_M(x, y) + \frac{M_{\text{mean}}}{2}. \tag{4}$$

where M_{\max} and M_{mean} are the maximum and average value of the image M , respectively, $B_M(x, y)$ is the value of bilateral filtering image of M at point (x, y) , $A_M(x, y)$ is the adaptation factor at point (x, y) , and $P(x, y)$ is the output image.

Step 2. The second local lighting compression imitates the Amacrine cells:

$$Q(x, y) = (P_{\max} + A_p(x, y)) \frac{P(x, y)}{P(x, y) + A_p(x, y)}. \tag{5}$$

$$A_p(x, y) = B_p(x, y) + \frac{P_{\text{mean}}}{2}. \tag{6}$$

where P_{\max} and P_{mean} are the maximum and average value of the image P achieved by the previous step, respectively, $B_p(x, y)$ is the value of bilateral filtering image of

P at point (x, y) , $A_p(x, y)$ is the adaptation factor at point (x, y) , and $Q(x, y)$ is the output image.

Step 3. The spatial band-pass filtering imitates the Inner Plexiform Layer: DoG filtering is a convenient method to achieve spatial band-pass behavior. Hence, we imitate the Inner Plexiform Layer by DoG filtering. It can be described as the following equation,

$$F = \left(\frac{1}{2\pi\sigma_L^2} e^{-\frac{x^2+y^2}{2\sigma_L^2}} - \frac{1}{2\pi\sigma_H^2} e^{-\frac{x^2+y^2}{2\sigma_H^2}} \right) * Q. \quad (7)$$

where σ_L and σ_H are the corresponding standard deviations of the low pass filters, respectively.

3 Experimental Results

In this section, two well-known illumination face databases (i.e., Yale B and CMU PIE) are used to evaluate the proposed method. PCA is employed to reduce dimension, and the Nearest Neighbor Classifier based on Euclidean distance is used for classification. In the phase of PCA, we reserved 99% energy of the principal components. The parameters σ_d , σ_s , σ_L and σ_H are set to 3, 0.1, 0.5 and 2 respectively. We compared our method with MSR [1], SQI [2] and Vu [4] under the same experimental conditions. The parameters of all compared methods can refer to their corresponding literatures.

3.1 Experiments on Yale B

Yale B [9] contains ten people under 64 different illumination conditions for nine poses. In our experiments, only frontal face images are chosen because we focus on dealing with varying illumination problem. All images are manually cropped by the positions of the eyes and resized to 96×84. The cropped images include only the face with as little hair and background as possible. According to the angle between the light source and the camera axis, these images can be divided into five subsets, i.e., subset 1 (0°-12°), subset 2 (13°-25°), subset 3 (26°-50°), subset 4 (51°-77°) and subset 5 (above 77°). In this experiment, Subset 1 is used for training, and other images from subset 2 to 5 are used for testing. The corresponding right recognition rates are shown in Table 1. From the table, one can observe that all results of the proposed method are higher than those of others.

Table 1. Recognition accuracy (%) of various methods on the Yale B

Methods	Subset 2	Subset 3	Subset 4	Subset 5
MSR [1]	100.00	98.33	97.14	93.16
SQI [2]	100.00	99.17	95.71	95.26
Vu [4]	100.00	99.17	93.57	93.68
The proposed	100.00	99.17	99.29	100.00

3.2 Experiments on CMU PIE

There are 68 persons with different pose, illumination and expression (PIE) in CMU PIE face database [10]. In this paper, we concern about how to eliminate the problem of varying illumination. The subset includes 21 different illumination conditions and contains 1428 face images (21×68). These images are all cropped in the same way as Yale B and resized to 64×64 . Fig. 1 shows 21 images of one person under different illumination and their corresponding outcomes obtained by our method. The previous 3, 4, 5 and 6 images were used for training, and the remaining 18, 17, 16 and 15 ones for testing, respectively. Experimental results listed in Table 2 show that the proposed method is superior to others.



Fig. 1. Original images and the corresponding outcomes

Table 2. Recognition rates (%) of various methods when choosing previous some images per person as training set

Training samples	MSR[1]	SQI[2]	Vu [4]	The proposed
3	75.08	92.16	88.32	97.22
4	84.26	94.38	93.77	98.88
5	95.13	97.89	97.98	98.81
6	95.01	98.14	97.65	99.41

4 Conclusions

In this paper, we propose an improved retina modeling for variable lighting face recognition. Bilateral filter was used to estimate the adaptation factor X_0 in the Naka-Rushton equation instead of Gaussian filter. This method not only preserves image details while getting image low-frequency information but also gets more accurate local lighting than Gaussian filtering does. Experimental results on two classical lighting databases show that the proposed method is robust to variable lighting.

Acknowledgments. This work is partially supported by Nanjing Institute of Technology Innovation Fund (Grant No CKJ2010009), the Natural Science Foundation of Shandong Province (Grant No ZR2011FM004), Technology Project of Education Department of Fujian Province (JA10226) and University Innovation Research and Training Program of Fujian Province (mjcx1006).

References

1. Jobson, D.J., Rahman, Z., Woodell, G.A.: A multiscale Retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* 6(7), 965–976 (1997)
2. Wang, H., Li, S.Z., Wang, Y.: Face recognition under varying lighting conditions using self quotient image. In: *Proc. of Conf. Automatic Face and Gesture Recognition*, pp. 819–824 (2004)
3. Meylan, L., Alleysson, D., Susstrunk, S.: Model of retinal local adaptation for the tone mapping of color filter array images. *J. Opt. Soc. Amer.* 24, 2807–2816 (2007)
4. Vu, N., Caplier, A.: Lighting robust face recognition using retina modeling. In: *Proc. IEEE Int'l Conf. Image Processing*, pp. 3289–3292 (2009)
5. Naka, K.-I., Rushton, W.A.H.: S-potentials from luminosity units in the retina of fish (cyprinidae). *Journal of Physiology* 185(3), 587–599 (2007)
6. Héroult, J., Durette, B.: Modeling Visual Perception for Image Processing. In: Sandoval, F., Prieto, A.G., Cabestany, J., Graña, M. (eds.) *IWANN 2007. LNCS*, vol. 4507, pp. 662–675. Springer, Heidelberg (2007)
7. Benoit, A., Caplier, A., Durette, B., Héroult, J.: Using Human Visual System modeling for bio-inspired low level image processing. *J. Comput. Vis. Image Understand.* 114(7), 758–773 (2010)
8. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: *Proc. IEEE Int'l Conf. Computer Vision*, pp. 839–846 (1998)
9. Georgiades, A.S., Belhumeur, P.N., Kriegman, D.J.: From few to many: Lighting cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intel.* 23(6), 643–660 (2001)
10. Sim, T., Baker, S., Bsat, M.: The CMU pose, lighting, and expression (PIE) database. In: *Proc. of Conf. Automatic Face and Gesture Recognition* (2002)

Automated Object Length Measurement Applied to AFM, STM and TEM Images Based on the Snake Model

Leandro Marturelli, Lilian Costa, and Geraldo Cidade

Universidade Federal do Rio de Janeiro - UFRJ,
Polo de Xerém, Brazil
marturelli@ufrj.br, lilian.t.costa@gmail.com,
gcidade@biof.ufrj.br

Abstract. This work describes the development of an automated method to measure the length of filament-shaped objects from Atomic Force Microscopy (AFM), Scanning Tunneling Microscopy (STM) and Transmission Electron Microscopy (TEM) images. The proposed methodology can determine the length of the object(s) of interest using image segmentation, where the Parametric Deformable Model (PDM) (especially the Snake Model - SM) was applied. The measurement procedure starts from the segmentation step, where an improved erosion is applied, resulting in a thin curve representation with only two endpoints (skeleton). The piecewise linear approximation concept was also employed to measure the total length of an object, based on the dimension of the pixel (unitary length) that composes the skeleton. The accuracy of the algorithm was evaluated using a high precision STM 7x7 silicon image and a sort of DNA filaments (AFM).

Keywords: Length Measurement, Snake Model, AFM, STM, TEM, DNA.

1 Introduction

The motivation for the development of a reliable algorithm capable to measure the length of objects automatically began with the DNA images obtained from the Atomic Force Microscopy (AFM) technique [1], having in view the importance of their nanometrological characteristics. Several efforts have been done so far in automating a method to determine DNA filament lengths. *Ficarra et al* [2] developed an automated algorithm to determine the DNA fragment size from AFM images and extract the molecular profiles. *Spisz* and co-workers [3] presented an automated sizing DNA program for segmentation using a threshold selection method, thinning and neighborhood pixel processing. Different software packages have been described and applied to DNA fragment length determination. One of them is the ImageJ[®] [4], that offers a manual procedure to trace segmented lines over the image in order to approximate the correct length of the object. This work proposes a technique that comprises image segmentation and length estimation of the object(s) of interest in an automatic way. The main idea is to segment the image using Parametric Deformable Models, especially the Snake Model (SM) [5], to better represent the edges of the object(s). The

improved erosion procedure was used to skeletonize the selected objects, and a linear piecewise approach was employed to estimate their lengths. The original proposal of the SM has been successfully applied in a variety of problems for computer vision and image analysis, such as motion tracking and segmentation. The SM consists basically on an elastic curve that can dynamically fits to the real object shape due to the action of internal (elastic) and external (derived from the image) forces. The simplest external force is a gradient of the image function, but another approach called Gradient Vector Flow (GVF) [6] allows the SM to push the curves into boundary concavities, converging in a more accurate representation of the borders. The GVF deformable model is a version of the SM that uses the GVF as an external force (GVF-SM); the GVF-SM is able to improve the precision of the segmentation, which is a determinant feature in the nanoscale. The proposed technique is not intended to replace any other methodology, but to show the efficiency of the GVF-SM approach over poor signal-to-noise ratio images.

2 Definitions

2.1 The Snake Model (SM)

The SM is a 2-D parametric deformable model that is defined as a curve $x(s) = [x(s), y(s)]$, $s \in [0, 1]$ - here assumed to be closed - that moves through the spatial domain of an image in order to minimize the energy functional $P = \int_0^1 \frac{1}{2} (\alpha |x'(s)|^2 + \beta |x''(s)|^2) + E_{ext}^i(x(s)) ds$ (1), where α and β are weighting parameters that control the Snake's tension and rigidity, respectively. The terms $x'(s)$ and $x''(s)$ denote the first and second derivatives of $x(s)$ with respect to s . The external potential function E_{ext}^i is derived from the image so that it takes on its smaller values at the features of interest, such as boundaries. Given a grey-level image $I(x, y)$, viewed as a function of continuous position variables (x, y) , typical external potential function (designed to lead a deformable contour toward step edges) is $E_{ext}^b(x, y) = -|\nabla G_\sigma(x, y) * I(x, y)|^2$ (2), where $G_\sigma(x, y)$ is a two-dimensional Gaussian and ∇ is the gradient operator. A deformable contour that minimizes P must satisfy the Euler equation $\alpha x''(s) - \beta x''''(s) - \nabla E_{ext} = 0$ (3).

The external potential force pulls the deformable contour toward the desired image edges, while the intrinsic forces minimize stretching and bending. Defining F_{int} as $\alpha x''(s) - \beta x''''(s)$ and F_{ext} as ∇E_{ext} , (3) can be rewritten as $F_{int} - F_{ext} = 0$ (4) representing a force balance equation. When the forces are balanced by means of (4), the state solution is achieved. A numerical solution can be found by discretizing (3) and solving the discrete system iteratively. The partial derivative of x with respect to t is then set equal to the left side of (3) $x_t(s, t) = \alpha x''(s, t) - \beta x''''(s, t) - \nabla E_{ext}$ (5).

2.2 The Gradient Vector Flow (GVF)

The GVF is another external potential function introduced by *Xu et al* [6] that can be used to improve the solution. The GVF is a vector diffusion method defined through the following diffusion-reaction equation

$$v_t = g(|\nabla E_{ext}|) \nabla^2 v - h(|\nabla E_{ext}|) (v - \nabla E_{ext}) \quad (6)$$

where $v_0 = \nabla E_{ext}$ and $g(\bullet)$

and $h(\bullet)$ are nonnegative weighting functions defined on the image domain, or can be set as constants. The relationship between weight parameters was observed by *Marturelli et al* [7] with respect to the quality of the generated field. Particular advantages over a SM are: i) insensitivity to initialization and ii) the ability to move into boundary concavities. The GVF-SM uses GVF as an external force: $x_t(s, t) = \alpha x''(s, t) - \beta x'''(s, y) + GVF$ (7).

2.3 Skeletonization

In this work, the skeleton is defined as an open curve that fits to a median line of a polygon. In order to compute the skeleton, the mathematical morphology erosion operator is applied in an improved way. First, a binary (black & white) image from the result of the GVF-SM is computed. From this image, the segmented object is eroded in order to wiredraw it. This procedure works repeatedly, analyzing each pixel belonging to the border of the object. In order to implement this procedure, a comparison between the analyzed pixel with a list of pre-defined 3x3 kernels (Fig. 1) was defined, following a simple rule: *The analyzed pixel will be removed if there exist two or more self connected neighbors. On the other hand, the pixel cannot be removed having in view that a segment fault may occur, resulting in disconnected curves for the same object.*

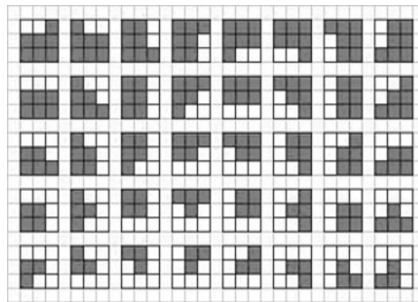


Fig. 1. Forty pre-defined kernels (templates) used to erode the object

2.4 Length Measurement

From the skeletons a piecewise linear approximation was used to measure the length of the object(s). The main idea is to compute the length unit of each pixel that forms

the skeleton, using 1 for vertical and horizontal directions and $\sqrt{2}$ for the diagonal. The sum of values results in a dimensionless length measurement that must be scaled to the original image dimension by means of multiplying the skeleton's length by a scale factor defined by the user.

2.5 Measurements Results

This section shows the results of some measurement experiments using AFM images of DNA filaments, in which an Intel[®] Core[™]2 CPU 6420 @ 2.13 GHz computer with 2 GB of RAM was used.

In order to evaluate the algorithm's capability to make precise measurements, a 7x7 silicon STM image was employed. This experiment was convenient and also essential due to the regular and precise characteristics of the silicon surface (Fig. 2(a)): the distance between two nano-holes belonging to a 7x7 cell corresponds to 2.68 nm. Fig. 2(b) illustrates the SM approach over Fig. 2(a), showing four measurements (A, B, C and D) with an average of 2.675 ± 0.01 nm.

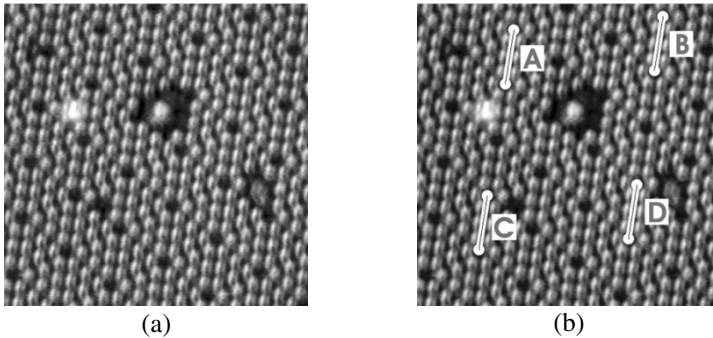


Fig. 2. (a) 512 x 512 STM image of 7x7 silicon cells (14 nm), and (b) the results after skeletonization (A=2.69nm, B=2.67nm, C=2.67 nm and D=2.67 nm)

In another experiment, an AFM image containing plasmid DNA filaments was used (Fig. 3(a)). When prepared, the sample was supposed to have tiny filaments with 896 base pairs each, which should represent separated filaments with 264 nm. A simple visual inspection shows that differences appear, due to, probably, different conformational arrangements after the sample preparation for AFM imaging. The results revealing such differences: filament 1 = 264 nm, 2 = 250 nm, 3 = 316 nm, 4 = 202 nm, 5 = 263 nm and 6 = 230 nm. Fig. 3(b) illustrates the skeletons drawn over the original image. In this case the method took 0.9 sec to process the measurements. The image was obtained from a Nanowizard AFM (JPK Instruments) using tapping mode in air.

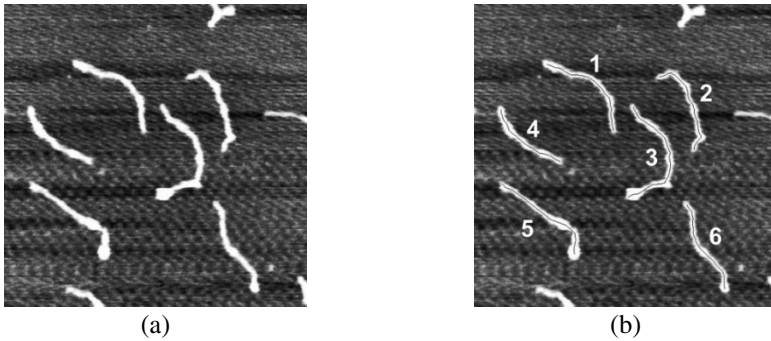


Fig. 3. (a) 512 x 512 AFM image of DNA filaments (700 nm x 700 nm) with 896 base pairs and (b) the skeletons after erosion (black lines)

The last experiment (Fig. 4), using an AFM image, shows the robustness of the algorithm with respect to its processing capacity and automation. First, two undesirable situations were observed: i) the presence of tiny objects that not represent a DNA filament and ii) the intersection of two or more DNA filaments forming one irregular object. In order to overcome these problems, two functionalities were introduced in the software implementation: one rejects objects that are smaller than a threshold and the other allows the user to discard the superimposed ones. Considering a huge amount of objects (approx. 100 DNA filaments) in the same image (Fig. 4(a)), the algorithm took 2 sec to process them, resulting in the skeletons shown in Fig. 4(b).

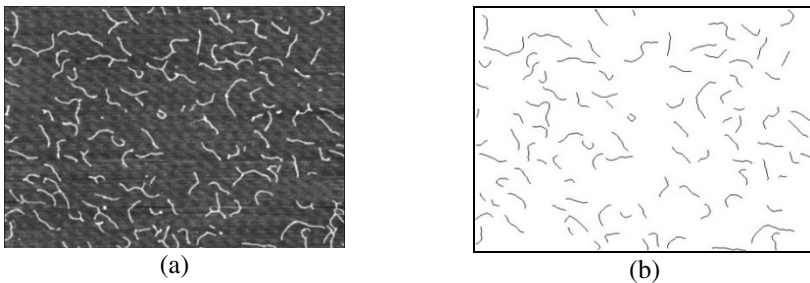


Fig. 4. (a) 512 x 387 AFM image of DNA filaments (3 μm x 2.27 μm) and (b) the skeletonization results

3 Conclusions

The methodology proposed in this work shows that the results of segmentation, supplied by GVF-SM, allow more accurate skeleton representation, which is an important feature to length measurement (the erosion procedure depends strongly on the initial shape of the object). Another characteristic explored on the GVF-SM was the ability to detect an amount of distinct objects in the same image, which contributes significantly to the automation of the process. A silicon surface image was used to evaluate

the algorithm's accuracy. This methodology is strong enough to be used in other applications, like, for instance, TEM images. Moreover, the time required to compute a great amount of objects was extremely short, increasing the efficiency of the methodology. For future work, the Topological Snake (T-Snake) [9,10] will be used to make the initialization easier, minimizing the inhomogeneous illumination problems.

Acknowledgments. The authors would like to thanks Ado Jório, Fernando Stavale, and Suzana Peripolli (researchers from Inmetro), for the images used in the experiments. This work was supported by CNPq, FINEP and FAPERJ.

References

- [1] Binnig, G., Quate, C.F., Gerber, C.: Atomic Force Microscope. *Phys. Rev. Lett.* 56(9), 930–933 (1986)
- [2] Ficarra, E., Benini, L., Macii, E., Zuccheri, G.: Automated DNA fragments recognition and sizing through AFM image processing. *IEEE Transactions on Information Technology in Biomedicine* 9(4), 508–517 (2005)
- [3] Spisz, T.S., Fang, Y., Reeves, R.H., Seymour, C.K., Bankman, I.N., Hoh, J.H.: Automated sizing of DNA fragments in atomic force microscope images. *Med. Biol. Eng. Comput.* 36, 667–672 (1998)
- [4] Rasband, W.: Image J. Software. National Institutes of Health (NIH)
- [5] Kass, M., Witkin, A.P., Terzopoulos, D.: Snakes: Active Contour Models. *Int. Journal. Computer Vis.* 1(4), 321–331 (1988)
- [6] Xu, C., Prince, J.: Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Proc.*, 359–369 (1998)
- [7] Marturelli, L.S., Giraldi, G.A., Rodrigues, P.S.S.: Gradient Vector Flow Models for Boundary Extraction in 2D Images. In: *Proceedings of Eighth IASTED - Computer Graphics and Imaging, Hawaii, USA (2005)* ISBN: 0-88986-512-4
- [8] Villarrubia, J.S.: Algorithms for Scanned Probe Microscope Image Simulation, Surface Reconstruction, and Tip Estimation. *J. Res. Nat. Inst. Stand. Technol.* 102, 425–454 (1997)
- [9] McInerney, T., Terzopoulos, D.: Topologically Adaptable Snakes. In: *Proc. of the Fifth Int. Conf. On Computer Vision*, pp. 840–845 (1995)
- [10] Marturelli, L.S., Giraldi, G.A., Rodrigues, P.S.S., Silva, R.R.: Improving the Initialization, Convergence and Memory Utilization for Deformable Models. In: *Handbook of Biomedical Image Analysis. Segmentation Models Part A, vol.1 (2005)* ISBN: 0306485508

Automatic Smudge Cell Recognition Based on Unimodal Region and Weak Edge Features

Guohui Qiao¹, Minglei Sun¹, Guanghua Zong¹, Fanggu Wu²,
Suling Huang³, and Shichuan Tang²

¹ School of Mechanical Engineering and Automation,
Beijing University of Aeronautics and Astronautics, Beijing 100191, China

² Beijing Municipal Institute of Labour Protection, Beijing 100054, China

³ Beijing Business School, Beijing 102209, China

kevin_joe@me.buaa.edu.cn

Abstract. Leukocyte differential count is a standard process in hematological diagnosing. The abnormal leukocytes impede the development of robust and efficient computer-assisted blood film analysis system, especially the smudge cells which are highly variable due to scrunching. This paper concentrates on the feature space construction for smudge cells. Unimodal regional features and weak edge characters are combined with histogram distribution statistics and gray level co-occurrence matrix to form the final feature space. A Supporting Vector Machine is adopted to construct the classifier. Experiment shows that the proposed features are efficient to compute and the sensitivity and specificity for smudge cell classification are both promising.

Keywords: smudge cell classification, feature extraction, unimodal distribution, weak edge features.

1 Introduction

Leukocytes are cells of the human immune system and differential counting of them provides vital important information for diagnosing of malignant blood diseases. Most of the blood samples could be analyzed automatically and efficiently through hematology analyzer. However, over 15 percent of blood samples require manual microscopic observation either because of biological rules or analyzer flags [1]. Manual microscopic examination of blood smears is not only time consuming and labor intensive but it also requires highly trained staff [2]. Automated blood film analysis systems have been introduced in the past and draw more and more attention [3-8]. Most of the proposed methods concentrate on the classification of five types of mature leukocytes (i.e., neutrophil, lymphocyte, monocyte, eosinophil and basophil). Smudge cells, which are mechanically distorted white cells, appear smudged in the blood smear and a significant number of smudges is a characteristic finding associate with chronic lymphocytic leukemia [9]. In the automated blood film analysis system, smudge cells still need to be reclassified by the clinician after pre-classification of the system. To reduce the requirement of reclassification for the operators, we research

on the feature extraction and differentiation of smudge cells from other types of leukocytes.

Little research has involved in the feature extraction of smudge cells. T. Bergen, etc [10] adopted a measure of unimodality proposed by Hartigan and Hartigan [11] to evaluate the distribution of cell histogram to distinguish smudge cells from other types of leukocytes. Experiment shows that the feature space of smudge cell needs to be expanded for a more precise classification rate. As the edges of smudge cells are usually blurred by scrunching and serves as an important feature in manual recognition, we quantified its weak edge features through the gradient information along the cell edge. In this paper, we combined the edge information with the unimodal region character and other texture features to form the feature space for smudge cell classification. Results show that smudge cells could be readily distinguished from other leukocytes, and the reclassification rate could be reduced for the automatic blood film analysis system.

2 Method

The proposed approach consists of multiple processes. The leukocyte and its component, plasma and nucleus, need to be segmented for further feature extraction. The feature space for smudge cell in this paper mainly consists of unimodal region features and weak edge features. We will detail each section in the following.

2.1 Leukocyte Segmentation

In a computer-assisted blood film analysis system, the detection and segmentation of leukocytes forms an essential step and its exactness is fundamental for the subsequent feature extraction and classification. Many research concentrates on this issue with an emphasis on Level-set method recently [12-15]. T. Bergen, etc [10] combines pixel-wise classification with template matching to locate erythrocytes and integrates their position information to revise the speed function calculated in Level-set for more precise segmentation. Leukocyte nucleus is then obtained by Otsu's adaptive thresholding [16] within the leukocyte region. Experiments show that their method performs well and is adopted as the segmentation scheme in this paper.

2.2 Unimodal Region Features

As the structure of smudge cells are destroyed, they fail to present clearly separated cytoplasm and nucleus. However, most of the mature leukocytes (except for basophil) contain cytoplasm and nucleus with clearly defined edge and different colors. Thus the histogram of mature leukocytes should exhibit bi-modal features and that of the smudge cells exhibit unimodal features. Experiment shows that cytoplasm and nucleus give the greatest difference in the saturation channel hence we conduct the histogram distribution analysis in this channel. Fig. 1 illustrates the saturation image for a

neutrophil and smudge cell and their corresponding histograms. The bi-modality of the histogram for the neutrophil presents obvious difference from the unimodality of histogram for smudge cell.

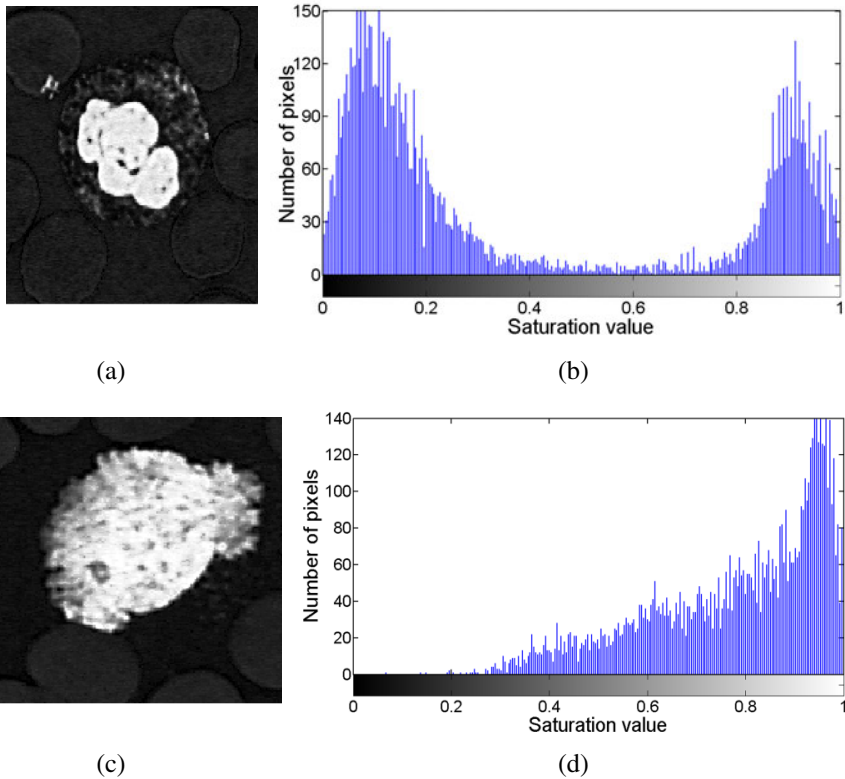


Fig. 1. Contrast between the histograms for the neutrophil and the smudge cell. (a) neutrophil; (b) corresponding histogram for the neutrophil; (c) smudge cell; (d) corresponding histogram for the smudge cell.

Hartigan and Hartigan [11] proposed the dip test which measures multimodality in a sample by the maximum difference. Compared with the differential method for multimodal analysis, the dip test is more robust and efficient. Thus we adopt this measurement in the modal analysis of histograms.

2.3 Weak Edge Features

Compared to the normal matured leukocyte, smudge cells exhibit low gradient on the edges, especially on the segmented nucleus edges. Figure 2 illustrates the computed gradient values on the nucleus edge of a neutrophil and the smudge cell in the saturation channel. The gradient of the smudge cell demonstrates more low values than the neutrophil on the edge.

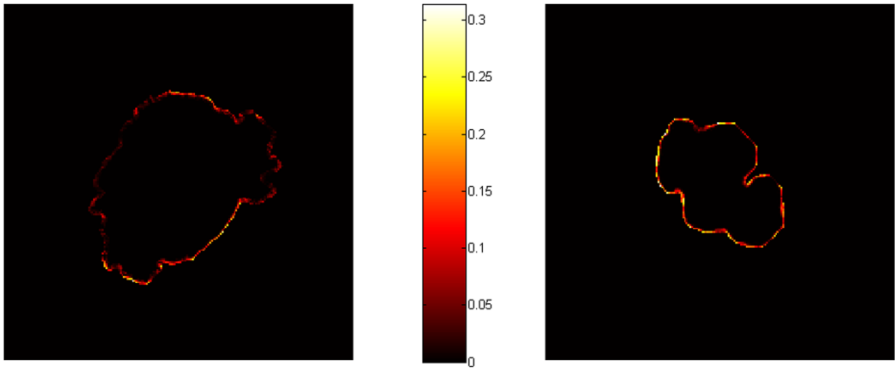


Fig. 2. Gradient difference for the smudge cell and mature leukocyte

Assume Nuc denotes the segmented nucleus region, and

$$grad(i, j) (0 \leq i \leq M - 1, 0 \leq j \leq N - 1, (i, j) \in \partial Nuc)$$

denotes the gradient of the nucleus edge in the saturation channel. We proposed the following features to quantify the edge status of the segmented nucleus:

(1) Average gradient value on the edge:

$$\bar{g} = mean(grad).$$

(2) Number of pixels under the thresholding gradient th

$$N = \#\{(i, j) \mid grad(i, j) < th, (i, j) \in \partial Nuc\},$$

where th is set 0.1 in this case.

The nucleus and cytoplasm are clearly separated for most of the normal mature leukocytes (except the basophil) whereas the contours of nucleus and cytoplasm are close for the smudge cells. We adopt a relative distance feature [17] RD to evaluate the distance between the contours of the nucleus and cytoplasm. Let n_1, n_2, \dots, n_e be the pixels on the contour of the nucleus, and c_1, c_2, \dots, c_e be the pixels on the contour of the cytoplasm. RD is defined as

$$RD = \frac{1}{2} \left(\sqrt{\frac{1}{n_e} \sum_{i=1}^{n_e} d_{ei}^2} + \sqrt{\frac{1}{c_e} \sum_{i=1}^{c_e} d_{ci}^2} \right)$$

where $d_{ei} = \min(dis(n_i, c_j) \mid j = 1, 2, \dots, c_e)$,

$d_{ci} = \min(dis(c_i, n_j) | j = 1, 2, \dots, n_e)$, and $dis(c_i, n_j)$ is the Euclidean distance between c_i and n_j .

3 Results

In order to evaluate the proposed features for automatic smudge cell and mature leukocyte differentiation, we selected 350 cell images (including 100 smudge cells and 250 mature leukocytes). Our blood films were stained with Wright Giemsa stain, and an Olympus BX51 microscope with a 100 \times oil immersion Plan Semi Apochromatic objective was utilized for image acquisition. A Supporting Vector Machine with a Radial Basis Function kernel was chosen as the classifier. The sensitivity for smudge cell is 91.0% and the specificity for mature leukocyte is 96.4%. Experiment shows that the proposed features are efficient to compute and perform well for smudge cells recognition.

4 Conclusion

This paper concentrates on the feature space construction for automatic smudge cell classification. Unimodal regional features and weak edge features are combined with histogram distribution statistics and gray level co-occurrence matrix to form the feature space. These features perform well for smudge cell and mature leukocyte differentiation thus could reduce the reclassification rate in an automated blood film analysis system. Future works should be focused on the automatic classification of abnormal leukocytes which are very helpful in clinical diagnosing, like blast cell, atypical lymphocyte and immature granulocyte.

Acknowledgment. This work was supported in part by the Beijing Municipal Natural Science Foundation No. 4092026, the National Natural Science Foundation of China (Grant No. 61071158), the Ph.D. Programs Foundation of Ministry of Education of China (Grant No. 20091102120022), and the Innovation Foundation of BUAA for PhD Graduates.

References

1. Cornet, E., Perol, J.P., Troussard, X.: Performance evaluation and relevance of the CellaVision (TM) DM96 system in routine analysis and in patients with malignant hematological diseases. *International Journal of Laboratory Hematology* 30, 536–542 (2008)
2. Briggs, C., Longair, I., Slavik, M., Thwaite, K., Mills, R., Thavaraja, V., Foster, A., Romanin, D., Machin, S.J.: Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system. *International Journal of Laboratory Hematology* 31, 48–60 (2009)
3. Rezaatofghi, S.H., Khaksari, K., Soltanian-Zadeh, H.: Automatic Recognition of Five Types of White Blood Cells in Peripheral Blood. In: Campilho, A., Kamel, M. (eds.) *ICIAR 2010. LNCS*, vol. 6112, pp. 161–172. Springer, Heidelberg (2010)

4. Osowski, S., Siroic, R., Markiewicz, T., Siwek, K.: Application of Support Vector Machine and Genetic Algorithm for Improved Blood Cell Recognition. *IEEE Transactions on Instrumentation and Measurement* 58, 2159–2168 (2009)
5. Nilufar, S., Ray, N., Zhang, H.: Automatic Blood Cell Classification based on Joint Histogram Based Feature and Bhattacharya Kernel. In: 42nd Asilomar Conference on Signals, Systems and Computers, vol. 1-4, pp. 1915–1918. IEEE, New York (2008)
6. Ramoser, H., Laurain, V., Bischof, H., Eckert, R.: Leukocyte segmentation and classification in blood-smear images. In: *Proceedings of the IEEE Engineering in Medicine and Biology Conference*, vol. 4, pp. 3371–3374 (2005)
7. Saeid Sanei, T.K.M.L.: Cell recognition based on PCA and bayesian classification. In: 4th International Symposium on Independent Component Analysis and Blind Signal Separation, Nara, Japan, pp. 239–243 (2003)
8. Ramoser, H.: Leukocyte segmentation and SVM classification in blood smear images. *Machine Graphics & Vision International Journal* 17, 187–200 (2008)
9. Gulati, G.: *Blood cell morphology: grading guide*. American Society for Clinical Pathology Press, Chicago (2009)
10. Bergen, T., Steckhan, D., Wittenberg, T., Zerfass, T.: Segmentation of leukocytes and erythrocytes in blood smear images. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1-8, pp. 3075–3078. IEEE, New York (2008)
11. Hartigan, J.A., Hartigan, P.M.: The dip test of unimodality. *Annals of Statistics* 13, 70–84 (1985)
12. Hamghalam, M., Motameni, M., Kelishomi, A.E.: Leukocyte segmentation in Giemsa-stained image of peripheral blood smears based on active contour. In: 2009 International Conference on Signal Processing Systems (2009)
13. Zhang, Y., Matuszewski, B.J., Shark, L.K., Moore, C.J.: Medical image segmentation using new hybrid level-set method. In: *Proceedings of Medivis 2008: Fifth International Conference Biomedical Visualization - Information Visualization in Medical and Biomedical Informatics*, pp. 71–76 (2008)
14. Eom, S., Kim, S., Shin, V., Ahn, B.-H.: Leukocyte Segmentation in Blood Smear Images Using Region-Based Active Contours. In: Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P. (eds.) *ACIVS 2006*. LNCS, vol. 4179, pp. 867–876. Springer, Heidelberg (2006)
15. Mukherjee, D.P., Ray, N., Acton, S.T.: Level set analysis for leukocyte detection and tracking. *IEEE Transactions on Image Processing* 13, 562–572 (2004)
16. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics SMC-9*, 62–66 (1979)
17. Sf, Y.-M., Yk, C., Yp, C.: Edge enhancement nucleus and cytoplasm contour detector of cervical smear images. *IEEE Transactions on Systems, Man and Cybernetics-PART B: Cybernetics* 38, 353–366 (2008)

Recognition of Touching Erythrocytes via Contour Radial Uniformity and Sector Region Distribution Features

Minglei Sun¹, Di Wang¹, Wen Wen¹, Rong Zhang¹,
Shichuan Tang², and Bin Zhang²

¹ School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, 100191 Beijing, China

² Beijing Municipal Institute of Labour Protection, 100054 Beijing, China
sunminglei@buaa.edu.cn

Abstract. The presentation of touching erythrocytes affects the Red Blood Cell counting and diagnosing in automatic blood cell analysis based on digital image processing, thus a robust and efficient recognition algorithm for touching erythrocytes is required. This paper analyses the typical features of touching erythrocytes and presents two new features to characterize them. The algorithm quantifies the radial uniformity of cell outer contours to describe the shape of the erythrocyte, and calculates the contour distribution within sector regions to efficiently approximate the gray scale distribution for texture representation. An ANN classifier was constructed based on the above features, and the discriminating rate for touching erythrocytes achieved 91.82%, which is promising and indicates the efficiency of the proposed features.

Keywords: touched erythrocytes, radial uniformity, sector areas, recognition.

1 Introduction

Complete Blood Count (CBC) is an important test conducted in the clinical laboratories that gives diagnosing information about the patient's blood. Erythrocytes, which make up the major blood cell components, often touch each other and bring counting and classification errors during automatic blood cell morphology analyzing. Hence a robust and efficient touching erythrocyte recognition algorithm based on digital image processing is urgently needed. As the touching erythrocytes are highly variant and irregular, recognition of them remains a difficult problem in image cytometry.

Recently many algorithms for touching erythrocytes recognition have been proposed[1-7], among which the general approaches are based on edge geometry[8] or the theories of morphology[9-11]. Experimental results indicate that discriminating limitations exist both in the concave point searching based on edge geometry[12] or watershed based on theories of morphology[13], because of the complexity of blood cell images. Thus current methods failed to meet the requirement of clinical diagnosis in automatic touching erythrocytes recognition. This paper summarizes the typical features of touching erythrocytes, and then proposes two new features to characterize

them, i.e., Contour Radial Uniformity (CRU) and Sector Region Distribution (SRD). Experiments showed that the classifier constructed based on the above features achieved a classification rate of 91.82% for touching erythrocytes. Results indicate that the proposed features are promising and efficient in touching erythrocyte recognition.

2 Feature Analysis

Touching erythrocytes are generally divided into 2 cases in morphology, which are “serial” and “parallel”. The “serial” refer to cells connected end to end, forming outer contours without any close area among cells. And the corresponding “parallel” refer that cells mutually surround into a closed region, forming both outer contours and inner contours. Fig. 1(a) and Fig. 1(b) show touching erythrocytes of the “serial” and the “parallel”.

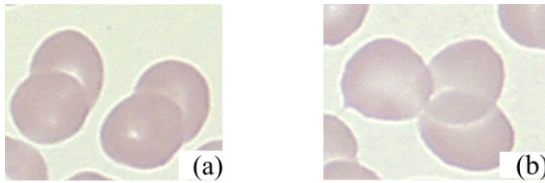


Fig. 1. The morphology of touching erythrocytes

Color features are not appropriate for touching erythrocytes recognition because both touching erythrocytes and single erythrocytes have the similar color set. Besides, different types of erythrocytes in blood cell images, such as macrocytes, have approximate size with touching erythrocytes. Therefore, geometric features, like area and perimeter, are not suitable for touching erythrocytes recognition.

This paper firstly analyses the manual criteria for recognition of touching erythrocytes. The human visual system determines that it is impossible to process too many features when recognizing touching erythrocytes. Usually, only one or two key features can successfully separate touching erythrocytes from single erythrocytes. Through analysis we know that the most salient features include concave point pairs on the outer contour and morphological differences between touching erythrocytes and single erythrocytes, as shown in Fig.2.

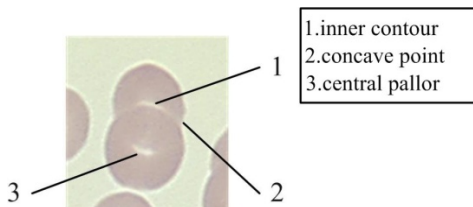


Fig. 2. The characteristic of touching erythrocytes

3 Feature Extraction

This paper proposes two new features to distinguish the touching erythrocytes from the single erythrocytes: Contour Radial Uniformity and Sector Region Distribution. We detail each feature as follows.

3.1 Page Numbering and Running Heads

A single erythrocyte shapes like a circle. The set of the cell radiuses is given by

$$U_{rs} = (r_1, r_2, r_3 \dots, r_N) \tag{1}$$

where N is the number of points on the contour, and r_1 to r_N are Euclidean distances from the geometric center to each contour point.

The outer contour of touching erythrocytes differs from the single RBC. We define Contour Radial Uniformity (CRU) to quantify the bias of a object shape towards a regular circle, which is given by

$$CRU = \frac{1}{N} \sum_{k=1}^N \left[\left\| (x_k, y_k) - (x_0, y_0) \right\| - \mu_R \right]^2 \tag{2}$$

where (x_k, y_k) is the coordinate of outer contour, and (x_0, y_0) is the geometric center of the object. μ_R denotes the average distance between the geometric center and the outer contour, given by

$$\mu_R = \frac{1}{N} \sum_{k=1}^N \left\| (x_k, y_k) - (x_0, y_0) \right\| \tag{3}$$

When a 2-D object varies towards a circle, CRU monotonically decreases and reaches 0. It could also be proved that CRU is translation, rotation and scaling invariant.

3.2 Caculation of Inner Sector Area Texture Features

As touching erythrocytes consist of different RBCs, their local texture features are similar with a single erythrocyte, while their global texture features have significant differences with that of a single erythrocyte. Most erythrocytes are round and have central pallors in them, while touching erythrocytes are erythrocytes overlapping together. The texture features are symmetrical verses the partition line, or the line connecting a pair of concave points.

Fig. 3(a) shows the simplified model of a single erythrocyte by simplifying its outer contour and central pallor contour into circular. Suppose there is a segment, with

one of its endpoints at the geometric center and the length of the segment equaling to the radius R . When the segment sweeps round the circle, the gray-scale of pixels in each sector should be approximately consistent.

Fig. 3(b) gives the simplified model of touching erythrocytes. Construct a circle with its center at the geometric center and of radius R_1 . The circle of radius R_1 has the same area with the touching erythrocytes area. So the sum of the pixels' gray-scale in each sector of the circle changes regularly. Two maxima and more minima, as well as gray-scale mutation will exist in theory.

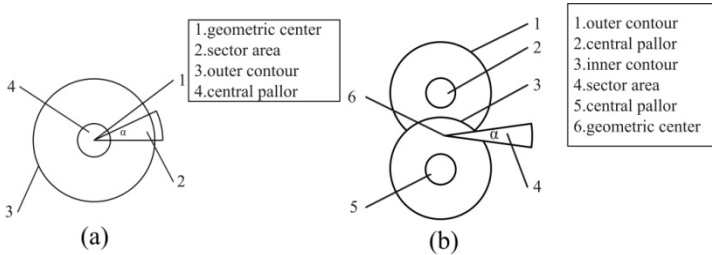


Fig. 3. The simplified model of a single erythrocyte and touching erythrocytes

In order to calculate the texture features of single erythrocytes and touching erythrocytes, we take the following steps:

1. Compute the geometric center O ;
2. Compute the radius R by $R = \sqrt{A/\pi}$, where A is the area of the cell region.
3. Calculate the gray-scale of pixels in each sector, with its center at O and of radius R and central angle α .

In the calculation of the feature of sector region, [14] proposed Upright Integral Image (UII), or the sums of gray-scale of pixels in each sector, defined as:

$$U(x, y) = \sum_{1 \leq x \leq n, 1 \leq y \leq m} I(x, y) \tag{4}$$

where $I(x, y)$ denotes gray-scale of pixel at (x, y) .

To simplify the computation of sector region textures, we propose the Sector Region Distribution (*SRD*) of contours obtained by canny operator to approximate the grayscale textures. Since we get canny edge directly from grayscale images, the value of pixels is either 0 or 1. The canny edge clearly shows texture differences between single erythrocytes and touching erythrocytes. Fig.4 (a) shows the canny edge of a single erythrocyte and Fig.4 (b) shows the canny edge of the touching erythrocytes.

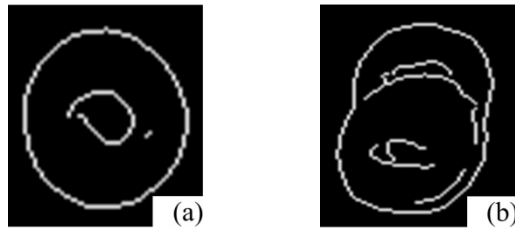


Fig. 4. The canny edges of: (a) single erythrocyte and (b) touching erythrocytes

The method to calculate SRD is as follows.

1. Calculate N_i , denoted by the number of pixels satisfying $f(x, y) = 1$ in each sector region, where i is the serial number of sector areas and $f(x, y)$ is the pixel value at (x, y) .

2. Compute SRD

$$SRD = \frac{1}{N} \sqrt{(N_i - \bar{N})^2} \quad (5)$$

where \bar{N} is the average of N_i .

4 Experimental Results

The algorithms of feature extraction and recognition in this paper are implemented in Matlab R2009b. 100 blood cell images were applied with it, each of them including about 40 erythrocytes.

4.1 Calculation of Radius Set

Fig.5 (a), Fig.5 (b), Fig.5 (c) respectively show the binary images of a single erythrocyte, touching erythrocytes comprising 2 RBCs and that comprising 3.

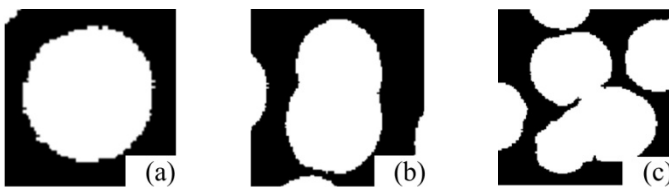


Fig. 5. The binary images of : (a) single erythrocyte and (b)(c) touching erythrocytes

The results of cell radiuses in these 3 cases are shown in Fig.6, where the abscissa represents serial numbers of boundary points and the ordinate represents cell radiuses. It can be seen from Fig.6 (a) that radius of a single erythrocyte varies within a small range, and none apparent feature points are found. In Fig.6 (b), there appear 2 minimum points, corresponding to the 2 concave points in Fig.5 (b), while the 3 minimum points in Fig.6 (c) correspond to the 3 concave points in Fig.5 (c). Since the concave point is an important feature for touching erythrocytes, it proves to be feasible to recognize touching erythrocytes though cell radiuses.

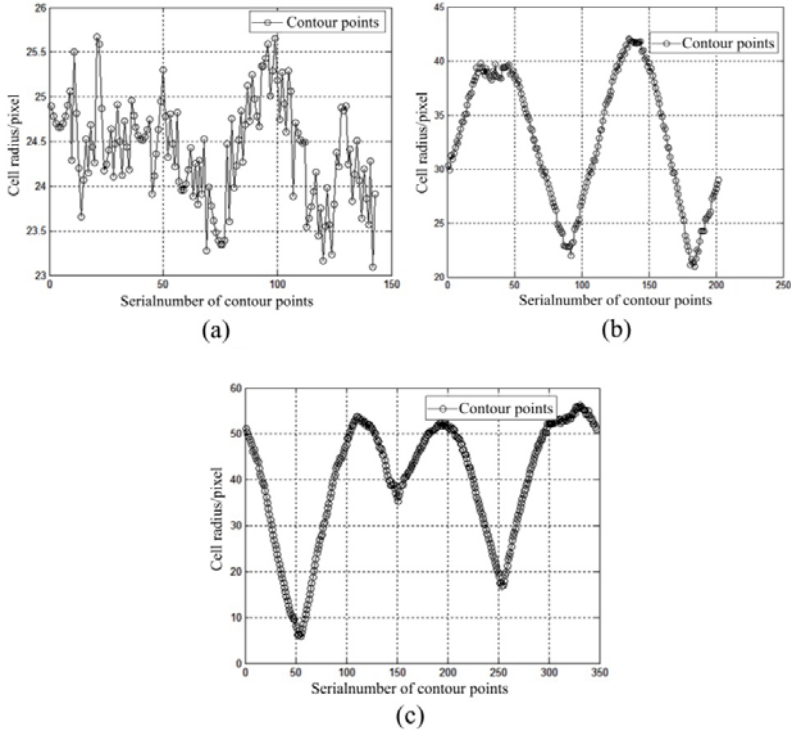


Fig. 6. The radius calculation of: (a) a single erythrocyte; (b) two erythrocytes touching; (c) three erythrocytes touching.

4.2 Calculation of CRU

Fig.7 shows the CRU calculated from 53 erythrocytes in a blood cell image. Among those the CRU value of 42 single erythrocytes are on the left side and that of 11 touching erythrocytes on the right side. The data shows significant differences between touching erythrocytes and single erythrocytes.

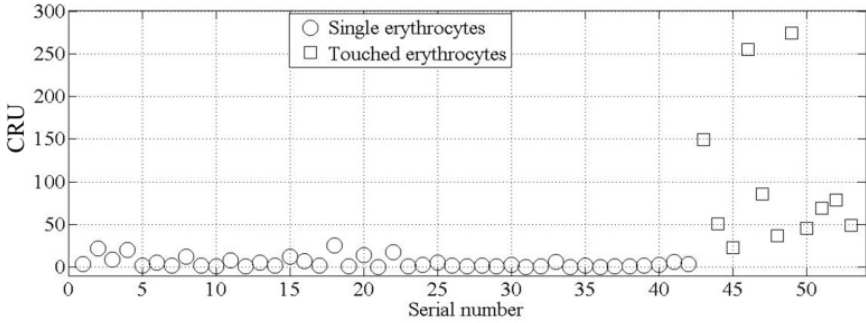


Fig. 7. The circularity of 53 erythrocyte regions in experiments

4.3 Calculation of SRD

Fig.8 shows SRD features of 53 erythrocyte regions in the above image. The experimental data indicates that touching erythrocytes can be distinguished from single erythrocytes by SRD. In addition, the extraction of canny edges directly influences the recognition because the statistical analysis is based on the canny edges in each sector region.

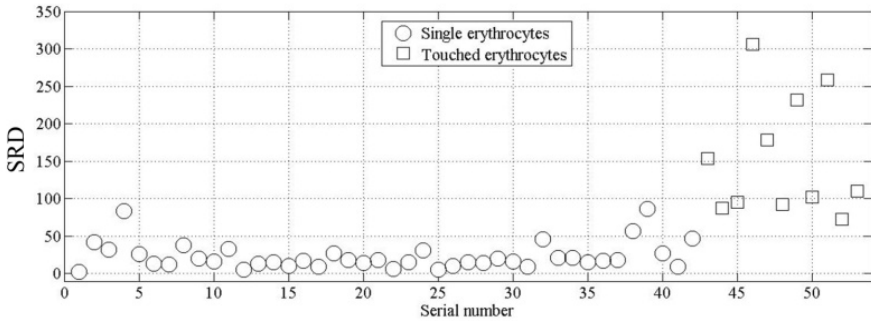


Fig. 8. Standard deviations of sums of pixels' gray scale in sector areas (53 erythrocyte regions in total)

We plot the 53 erythrocytes by combining CRU and SRD features and obtained in Fig.9. The figure shows that among the total 53 erythrocytes, 11 touching erythrocytes regions and 42 single erythrocytes regions are linearly separable. Besides, misclassification could be avoided in the 53 erythrocytes by combining the above features.

The linear equation of the classification boundary can be obtained by using neural network, since touching erythrocytes and erythrocytes regions are linearly separable. This paper adopts one-neuron perceptron. The experimental data are 100 images collected from the blood smear analyzer. 20 of them are for training the neural network and the others are for test. The experimental result shows the recognition rate is up to 91.82%.

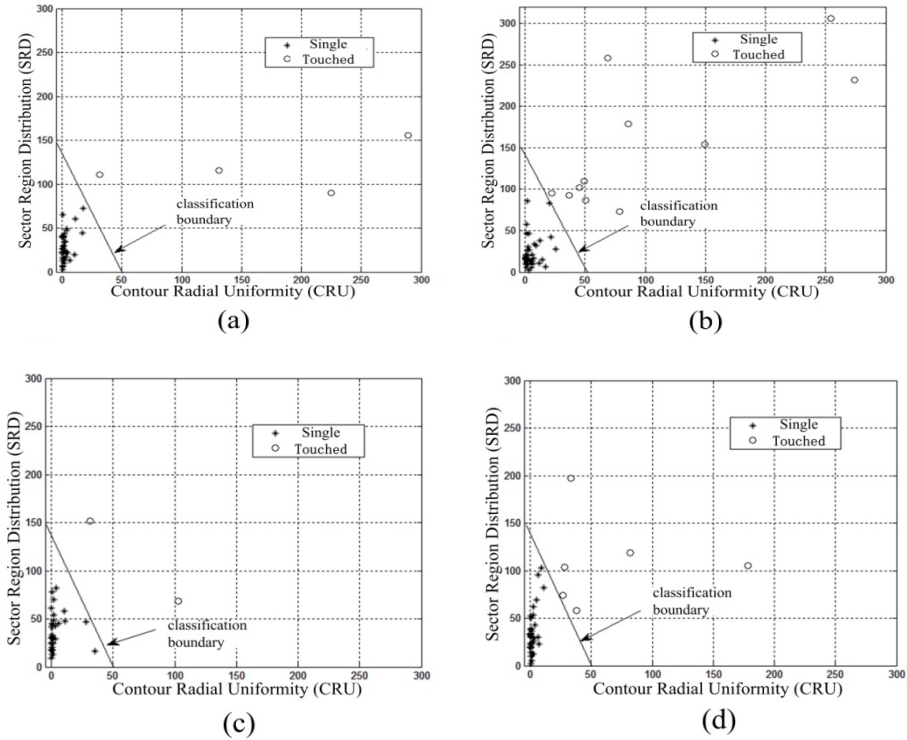


Fig. 9. The recognition result of touching erythrocytes

5 Conclusion

This paper proposed two new features, Contour Radial Uniformity and Sector Region Distribution, to discriminate the single erythrocyte from touching erythrocytes and provides the method to for calculation. Theory analysis and experimental results show its effectiveness to combine the two features for the touching erythrocytes recognition. Therefore, the algorithm is useful in the field of automatic blood cell morphology analysis.

Acknowledgments. The authors would like to thank their colleagues Guohui Qiao and Weilong Liu in the Robotics Institute of BUAA, Prof. Guanghua Zong for their advice on the configuration of the field. This research was supported in part by the Beijing Municipal Natural Science Foundation No. 4092026, the National Natural Science Foundation of China (Grant No. 61071158), and the Ph.D. Programs Foundation of Ministry of Education of China (Grant No. 20091102120022).

References

1. Fu, R., Shen, H.: Automatically discriminating overlapped cell based on shape factor analysis: a technical study. *Bulletin of The Academy Of Military Medical Sciences* 31, 463–497 (2007)
2. You, Y., Yu, H.: A Separating Algorithm for Overlapping Red Blood Cell Images. *Journal of Image and Graphics* 10, 736–740 (2005)
3. Lu, J., Yang, J., Tang, Z., Ye, Y., Wang, D.: Design of A Separating Algorithm for Overlapping Cell Images. *Journal of Computer Research and Development* 37, 228–232 (2000)
4. Yang, Z., Tao, Q., He, X.: Cell image separation based on shortest distance. *Journal of Chengdu University of Information Technology*, 377–380 (2004)
5. Nilsson, B., Heyden, A.: Model-based segmentation of leukocytes clusters. In: *Proc. 16th Int. Conf. Pattern Recog., Quebec, QC, Canada, vol. 1*, pp. 727–730 (2002)
6. Casasent, D., Talukder, A., Cox, W., et al.: Detection and segmentation of multiple touching product inspection items. In: *Proc of Int. Conf of SPIE, Boston, vol. 2907*, pp. 205–215 (1996)
7. Hero, H., Susanne, S., Madhukar, P.: Analysis of blood and bone marrow smears using digital image processing techniques. In: *Proceedings of the SPIE Medical Imaging, San Diego, Edb*, pp. 624–635 (2002)
8. Lu, Z., Tong, T.: The Application of Chain Code Sum in the Edge form Analysis. *Journal of Image and Graphics* 7, 1323–1328 (2002)
9. Zhang, J., Hu, P.: A Survey for the Arithmetic of Overlapped Cell Segmentation. *Computer & Digital Engineering* 36, 53–81 (2008)
10. Lu, Z., Fan, Y., Pang, Q.: Comparison of Several Algorithms Based on Mathematical Morphology of Overlapped Cells Segmentation. *Computer Engineering and Applications* 40, 57–59 (2004)
11. Haralick, R.M., Stenberg, S.R., Zhuang, X.: Image analysis using mathematical morphology. *IEEE Trans. PAMI*, 532–533 (1987)
12. Fu, R., Shen, H., Chen, H.: Research of automatically separating algorithm for overlap cell based on searching concave spot. *Computer Engineering and Applications* 43, 21–23 (2007)
13. Ma, D., Cao, P., Pan, K., Cheng, J.: Comparison of Some Methods for Segmentation of Overlapped Nuclei. *Beijing Biomedical Engineering* 28, 142–147 (1999)
14. Chen, L., Chen, B., Chen, J.: Features and application of image sector. *Journal of Computer Applications* 28, 2896–2899 (2008)

A Continuation Log-Barrier Method for ℓ_1 -regularized Least Square

Min Zhang¹ and Dongfang Chen²

¹ College of Computer Science & Technology,
Wuhan University of Science & Technology, Wuhan, China
xtt654@sina.com

² College of Computer Science & Technology,
Wuhan University of Science & Technology, Wuhan, China
df_chen@163.com

Abstract. Recently, there are increasing attention paid on compressed sensing which is distinct different from traditional signal processing and image reconstructed from indirect or incomplete measurement, especially ℓ_1 -norm problem which is transformed from compressive sensing. The idea of ℓ_1 -regularization, as the one of ℓ_1 -norm, has been receiving a lot of interest in signal processing, image recovery and statistic, etc. This paper will introduces a continuation log-barrier method for solving ℓ_1 -regularized least squares problem in the field of compressive sensing, which is a second-order method. Our work is inspired by the work in [4] and continuation idea, and the paper will introduce the continuation technique to increase the convergence rate. Therefore, Our continuation log-barrier method for ℓ_1 -regularized least square problem is accurate and fast in the sense.

Keywords: compressed sensing, ℓ_1 -regularization, least squares, log-barrier, continuation, convex optimization.

1 Introduction

Compressed Sensing(CS) attracts more people's attention in recent years .Then exactly what is compressed sensing? The so-called compressed sensing, the core of the concept is trying to reduce the cost of a signal or image measurement in the principle. That is to say , CS is assigned to the idea of encoding a large sparse signal using a relatively small number of linear measurements, and minimizing the ℓ_1 -norm (or its variants) in order to decode the signal. Similarly, CS shows that original signals or images can be reconstructed from fewer data/measurement than what is usually consider necessary. For more information on compressed sensing, see, for example, [1, 2,5,17,22,10].

In brief, CS demonstrates that solving an under-determined linear system of equations $Ax=b$ in mathematics in essence. If the "observation" b is contaminated with noise, then an appropriate norm of the residual $Ax-b$ should be minimized or constrained. Such considerations generate a family of related optimization problems. So the ℓ_1 -regularized least squares problem would be appropriate.

As is known to all , we demand high accuracy, low computation complexity, both accurate and fast algorithm to handle noise signal or image. As far as I'm concerned, [13] introduces a fixed point continuation technique which has been used successfully to increase the speed of convergence, that is, reduce the running time, the result is better when dealing with large scale problem.

This paper build on the work of S.-J. Kim, K. Koh, Boyd, D. Gorinevsky [4] , Candes, Romberg [23], and continuation strategy[24] , and proposes an improved algorithm for solving ℓ_1 -regularized least squares problems from incomplete measurement. And the paper is organized as follows. In section 2, the paper gives a brief but the essence of the summary of Boyd's work[4] and continuation technique. Then we demonstrate our improved log-barrier method with continuation in Section 3. In section 4, we present several experiments and analysis. Finally, we summarize the computational complexity and accuracy of our method.

2 Boyd's Method and Continuation Strategy

2.1 Boyd's Method

In the work [4],

$$\min_{x \in \mathbb{R}^n} \lambda \|x\|_{\ell_1} + \|Ax - b\|_{\ell_2}^2 \tag{1}$$

Boyd, Kim and his colleagues have shown that the solution x of (1), the form (1) firstly is transformed into Lagrange dual suboptimal problem,

$$\text{minimize } z^T z + \lambda \|x\|_1, \text{ s.t. } z = Ax - b$$

the Lagrangian is $L(x, z, v) = z^T z + \lambda \|x\|_1 + v^T (Ax - b - z)$

Secondly, the form(1) is transformed into a convex quadratic problem

$$\min \|Ax - b\|_{\ell_2}^2 + \lambda \|u\|_1, \text{ s.t. } -u \leq x \leq u$$

where ,the initial value of x set 0, the initial value of u set 1

they use log-barrier method,

$$f(x, u) = t \|Ax - b\|_{\ell_2}^2 + t\lambda \|u\|_1 - \sum_i^n \log(-f1) - \sum_i^n \log(-f2)$$

where, $f1 = x - u$, $f2 = -x - u$

The method takes the flow chart below:

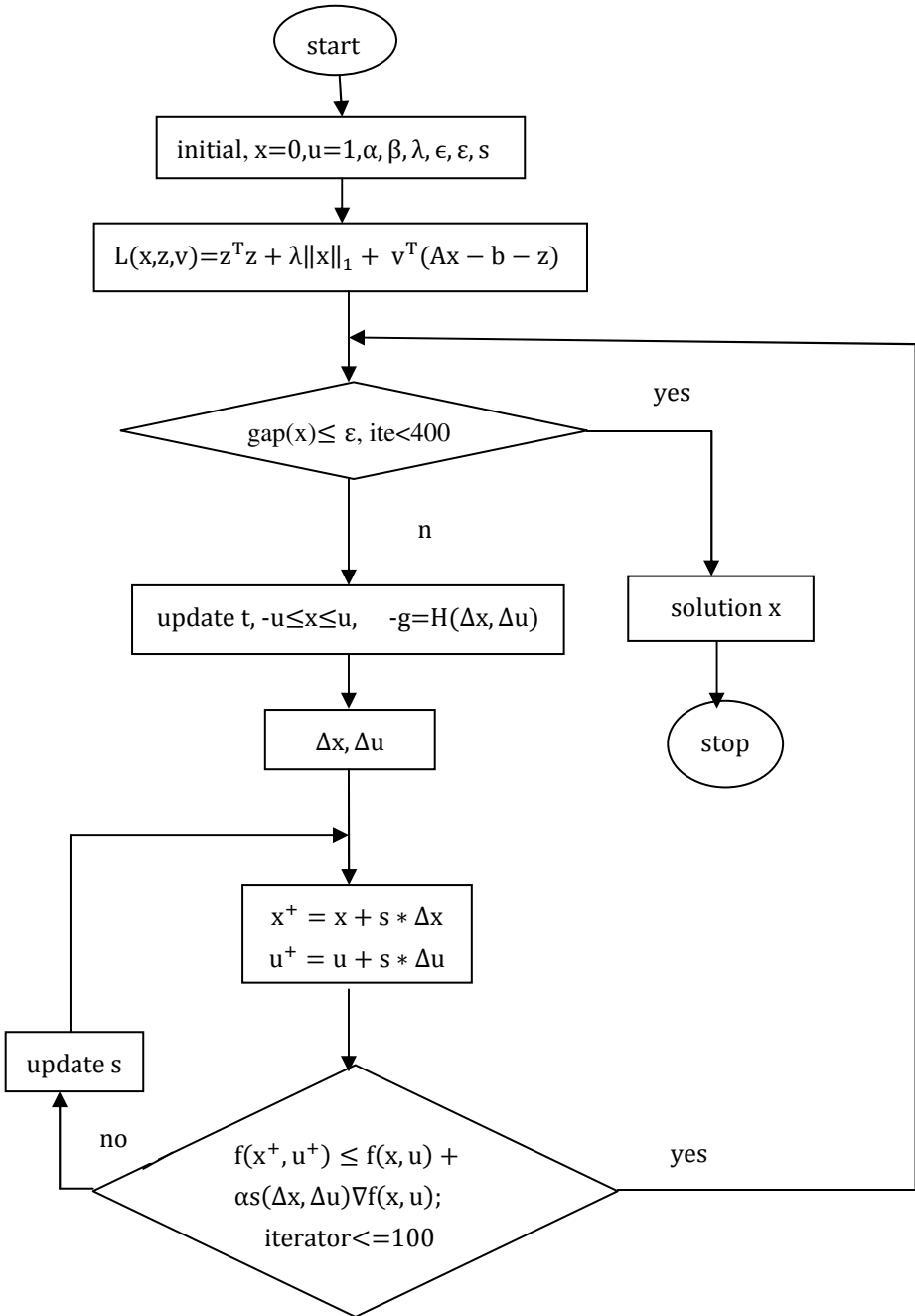


Fig. 1. The flow chart of work[4]

2.2 Continuation Strategy

In the following form,

$$\text{minimize } \lambda \|x\|_{\ell_1} + \frac{1}{2} \|b - Ax\|_{\ell_2}^2$$

for a fixed λ_f , the idea of continuation is simple: put forward a series of problems with reducing the value of argument λ , eg., $\lambda_0 > \lambda_1 > \dots > \lambda_f$, and use the intermediate solution as a warm start for the next problem. The continuation technique has been proved to be a very successful tool to increase the rate of convergence[13], especially when dealing with large-scale problems and high dynamic range signals.

3 Continuation Log-Barrier Method

3.1 General Formulation

In compressed sensing, one acquire a signal $x \in \mathbb{R}^n$ by sampling form

$$y = Ax + e, \tag{2}$$

where x is a signal of unknowns, $y \in \mathbb{R}^m$ is the observation of signal x ($m < n$). $A \in \mathbb{R}^{m \times n}$ is the sensing matrix. Moreover, e is a small noise, since there always exist noise in the real world.

It is hard to get solution x from (2) because it's under-determined equation when $m < n$ [21] and taking no account of e . To reconstruct x , one original approach in CS attempts to solving

$$\min \|x\|_{\ell_0}, \text{ s.t. } Ax = b \tag{3}$$

for (3) linear equation, since ℓ_0 -norm is intractable, under suitable conditions, minimizing the ℓ_1 -norm is equivalent to minimizing the so-called " ℓ_0 -norm" which is the number of nonzero, and Candes and Tao have proved that the ℓ_1 -norm approach is equivalent [21]

$$\min \|x\|_{\ell_1}, \text{ s.t. } Ax = b \tag{4}$$

So far, a lot of research and recovery algorithm is based on the (4) which is often preferred.

Since there always is noise in b in the case of real-world compressible signals, the other frequently approach consider to solve

$$\min \|x\|_{\ell_1}, \quad \text{s.t. } Ax - b \leq \delta$$

Later, for convenience of notation and computation, it is reformulated followed

$$\min \lambda \|x\|_{\ell_1} + \|Ax - y\|_{\ell_2}^2 \tag{5}$$

where $\|x\|_{\ell_1} = \sum_i^n |x_i|$ denotes the ℓ_1 -norm of x and $\lambda > 0$ is the regularization parameter, which is called ℓ_1 -regularized least squares program and has the analytic solution.

3.2 Log-Barrier Method

Before measurement ,for sake of reconstruction exactly, we should take full advantage of available fast transforms like FFT and DCT to make image sparse.

In order to solve the (5) expediently, the form (5) can be transformed to a convex quadratic problem with inequality constraints:

$$\min \lambda \|x\|_{\ell_1} + \|Ax - y\|_{\ell_2}^2, \text{ subject to } -u \leq x \leq u, \|Ax - y\|_{\ell_2}^2 \leq \varepsilon^2 \quad (6)$$

The standard log-barrier method transforms (6) into a series of nonlinearly constrained programs, the objective function:

$$f = \lambda \|x\|_{\ell_1} + \|Ax - y\|_{\ell_2}^2 + \phi(x, u) \quad (7)$$

where f is the objective function,

$$\phi(x, u) = -\left(\frac{1}{t}\right) (\sum_i \log(-f_1(x, u)) + \sum_i \log(-f_2(x, u)) + \log(-f_\varepsilon)),$$

$$\text{and } f_1(x, u) = x - u, \quad f_2(x, u) = -x - u, \quad f_\varepsilon = \left(\frac{1}{2}\right) * (\|Ax - y\|_{\ell_2}^2 - \varepsilon^2)$$

where, we often take the constant ε a little bigger than noise.

According to the convex optimality condition ,when $\nabla f = 0$ as f decrease, f will has the minimum, and at the current iterator x is the optimal solution.

Given that a point (x ,u) satisfied

$$\nabla f(x + \Delta x, u + \Delta u) = \nabla f(x, u) + \nabla^2 f(x, u)(\Delta x, \Delta u)$$

the point (x, u) is the optimal solution at the current iterator .

In primal barrier method[3], Newton’s method is used to minimize f ,i.e., the search direction $(\Delta x, \Delta u)$ is computed as the exact solution to the Newton system

$$tH \begin{bmatrix} \Delta x \\ \Delta u \end{bmatrix} = -tg \quad (8)$$

where H is the Hessian matrix and $H = \nabla^2 f \in R^{2n \times 2n}$, $g = \nabla f \in R^{2n}$ is the gradient at the current iterator (x ,u).

For faster convergence, however, we can update the parameter t at each iteration.

The following table shows the difference of the two method:

Table 1. The difference between the work[4] and our method

	Initial value	algorithm	Stop criterion
The work[4]	$X_0=0, u_0=1$	$Z=Ax-y; -u \leq x \leq u$	$Gap(x) \leq \epsilon, \text{ iterator} < 400$
Our method	$X_0=A^*y, u_0 \approx X_0$	$-u \leq x \leq u, \ Ax - y\ _2^2 \leq \epsilon$	$\nabla f * (\Delta x, \Delta u) \leq \epsilon, \text{ iterator} < 50$

This lead to the following flow chart:

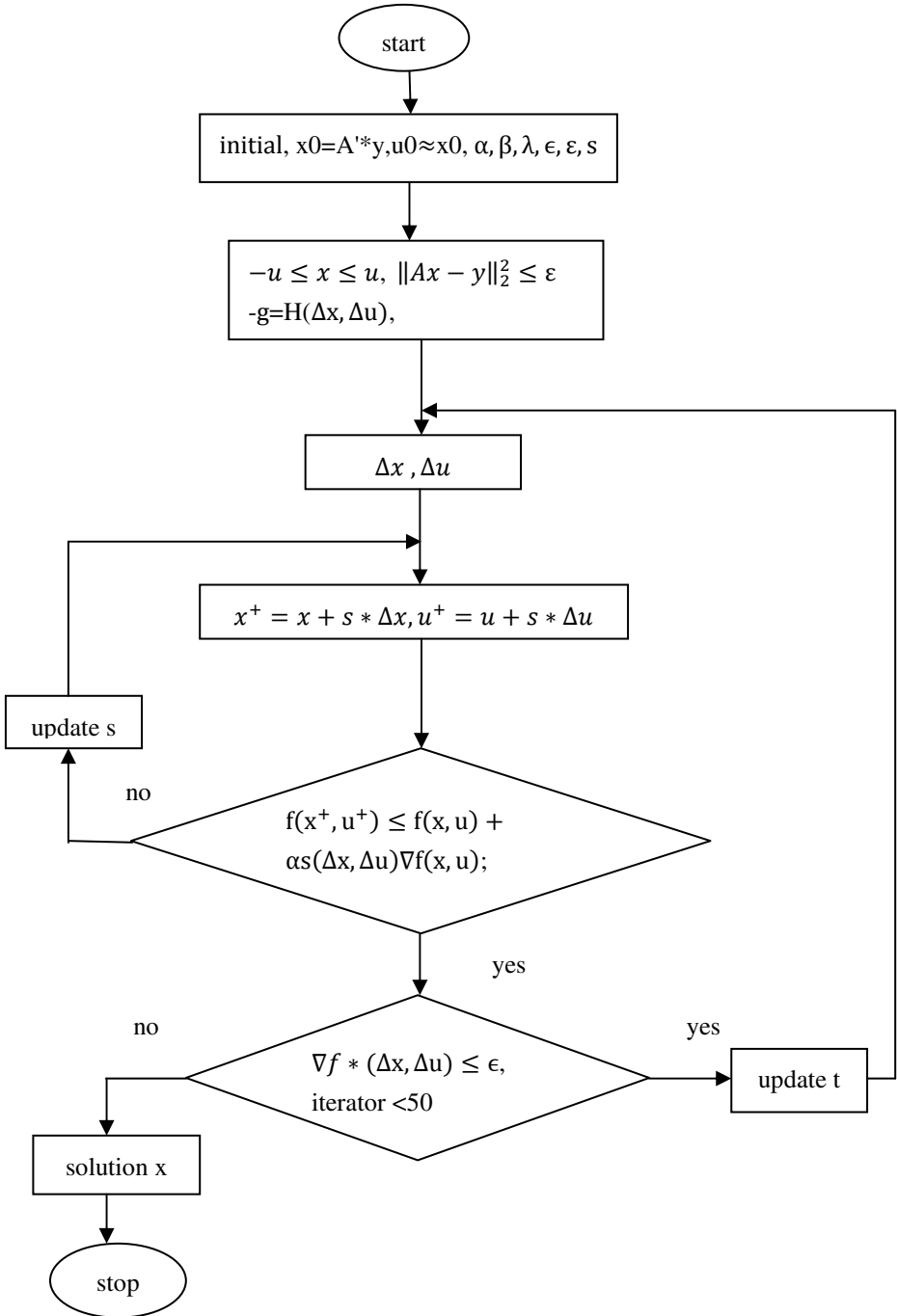


Fig. 2. The flow chart of our method

3.3 Continuation

Recently, continuation widely applied to all kinds of optimization algorithm. Instead of solving problem (5) directly from scratch , we use a continuation strategy to solve a sequence of problem. Experiments in [10] have further confirmed the effectiveness of continuation. Therefore, we embed continuation procedure in our method.

In the form(7), the continuation present that decreasing values of the parameter t , $t_0 > t_1 > \dots > t_f$, and use the intermediate solution as a warm start for the next problem. We will computer the value of the function f using the new the value of t , to speed up the convergence.

3.4 Search Direction

In form (7), the search direction should be obtained. We will get the search direction by the equality of the form (8).

In the form (8), We can find compact representations of the Hessian and gradient . The Hessian can be written as

$$H = \nabla^2 \|Ax - y\|_2^2 + \nabla^2 \phi(x, u) = \begin{bmatrix} 2 * A^T * A & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}$$

where, we set $r = Ax - y$, $A_t = A^T$, $atr = A_t * r$, and

$$\begin{aligned} \Sigma_{11} &= \Sigma_{22} - (1/f_\epsilon) * A_t * A + (1/f_\epsilon)^2 * atr * atr^T; \\ \Sigma_{12} &= -1./f1.^2 + 1./f2.^2; \\ \Sigma_{22} &= 1./f1.^2 + 1./f2.^2; \end{aligned}$$

The Hessian is positive definite. The gradient can be written as

$$g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \in R^{2n}$$

where $g_1 = \nabla_x f = 2 * A_t * A - \frac{1}{f_1} + \frac{1}{f_2} - \frac{1}{f_\epsilon} * A_t * (Ax - y)$

$$g_2 = \nabla_u f = \lambda * 1 + \frac{1}{f_1} + \frac{1}{f_2}$$

We compute the search direction approximately, applying the PCG algorithm [16, Sect. 5] to the Newton system (8). The PCG algorithm uses a preconditioner $P \in R^{2n \times 2n}$, which is symmetric and positive definite.

$$P = \begin{bmatrix} 2\tau I & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Sigma_{22} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix} + \begin{bmatrix} -(1/f_\epsilon) * A_t * A + (1/f_\epsilon)^2 * atr * atr^T & 0 \\ 0 & 0 \end{bmatrix}$$

The PCG algorithm needs a good initial search direction and an effective truncation rule. Here, initial point is 0. The truncation rule is exceed our define $pcgtol=0.1$.

4 Numerical Experiments

In the subsection, we give some numerical experiments to demonstrate the performance of the continuation log-barrier method by using 8G main memory and 8 nuclear of CPU. The algorithm parameters are taken as

$$\begin{aligned} \lambda=0.01, \quad \beta=0.5, \quad s_{\min}=0.5, \quad \alpha=0.01 \\ \mu=2, \quad \varepsilon=0.01, \quad N_{\text{pcg}}=200 \end{aligned}$$

Assume that, x is sparse or compressive and A is a matrix that satisfy Restricted Isometry Property(RIP)[1] such as Gaussian sample matrix, and recovery x from y by solving ℓ_1 -regularized least squares minimization problems. Hence, we use the following data. The number of signal x is 1024, the number of measurements matrix is 256×1024 , and the number of nonzero of x is about 51. The value of the nonzero components are generates an integer random value. And the stopping criterion is

$$|f_i - f| \leq 1e - 3 \quad \text{and} \quad \nabla f * (\Delta x, \Delta u) \leq \epsilon$$

The recovery result of 1D signal from incomplete measurement is as follows:

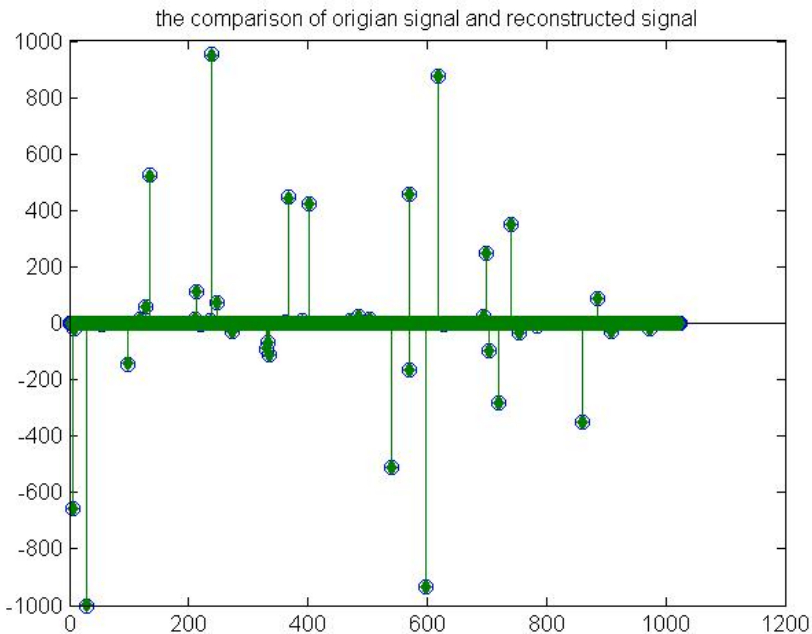


Fig. 3. The accurate recovery signal by our log-barrier method; solid line: original signal. Dots-o: recovered signal. In the figure, recovered signal is accurate

We apply our method to solve a 2D image recovery problem from incomplete frequency measurement. Here, we use the 256×256 classical Lena, and the result is as follows.



Fig. 4. Left: Original 256×256 image. Right: accurate recovery image by our method for ℓ_1 sp.

However, the work [4] has the performance of the following figure with the same 256×256 image:

reconstructed: psnr=25.4152, runtime(s)=583.02



Table 2. The results of two method

	Runtime(s)	psnr
The work[4]	583.02	25.4152
Our method	62.586	32.4613

Fig. 5. The recovery image by the work[4]

From the table above, we can see that our method is need less time and better accuracy when solving the same problem of ℓ_1 -regularized least square . In our work ,it needs less iteration on theory analysis, the computational complex of our method is $O(nm^{1/2})$ and the computational complex of [4] is $O(nm)$, which is computationally efficient. Results of the compared show that our method has improve the overall speed of convergence and save time.

Our method spend less time than [4],most of the reason is the original value of x and u . In [4],they set $x^0=0,u^0=1$,this need more time to approach to u for x , consequently lead to the increase of iteration and the computational effort. And we introduce the

continuation strategy to reduce the running time greatly. Moreover, the time will increased rapidly with the signal and image reconstruction of the increase of dimension.

5 Conclusion

In short, continuation technique will give a good tradeoff of computational effort versus the convergence rate it provides in our method. And it increases the speed of convergence with continuation. Therefore, Our method is far faster than Boyd[4] and other similar methods under the same conditions including restoration effect and the measurement .

In this paper, we improved an efficient method to solve (5). Focusing on the problem of ℓ_1 -regularized least square minimization, we show that the continuation log-barrier method is an efficient optimization technique to solve high-dimension signal. In our method, we apply the improved log-barrier convex method, to converge on the optimal solution and receive the good result. And numerical experiments indicate that our method is very efficient in ℓ_1 -regularized least square for compressed sensing field.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (Grant Nos. 61100055 and 60970065), ZJNSF under grant No.R1110261.

References

- [1] Candès, E.: Compressive sampling. In: Proc. Int. Congr. Mathematics (2006)
- [2] Donoho, D.: Compressed Sensing. IEEE Trans. Inf. Theory 52(4), 1289–1306 (2006)
- [3] Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)
- [4] Kim, S.-J., Koh, K., Lustig, M., Boyd, S., Gorinevsky, D.: A method for large-scale ℓ_1 -regularized least squares. IEEE Journal on Selected Topics in Signal Processing 1(4), 606–617 (2007)
- [5] Candès, E., Romberg, J.: Quantitative robust uncertainty principles and optimally sparse decompositions. To appear in Foundations of Comput. Math. (2005)
- [6] Mark Schmidt. Least Squares Optimization with L_1 -Norm Regularization. CS542B project Report (December 2005)
- [7] Wright, S.J.: Primal-Dual Interior-Point Methods. SIAM Publications (1997)
- [8] Nesterov, Y.E., Nemirovski, A.S.: Interior Point Polynomial Methods in Convex Programming. SIAM Publications, Philadelphia (1994)
- [9] Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. Ser. B 95, 3–51 (2003)
- [10] Wright, S.J., Nowak, R.D., Figueiredo, M.A.T.: Sparse reconstruction by separable approximation. IEEE Trans. Signal Proc. 57, 2479–2493 (2009)
- [11] Candès, E., Romberg, J.: ℓ_1 -Magic: Recovery of Sparse Signals via Convex Programming (2005)
- [12] <http://www.stanford.edu/>
- [13] Hale, E.T., Yin, W., Zhang, Y.: A Fixed-Point Continuation Method for ℓ_1 -Regularized Minimization with Applications to Compressed Sensing. CAAM Technical Report (2007)

- [14] Candès, E., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. *Comm. on Pure and Applied Math.* 59(8), 1207–1223 (2006)
- [15] Van Den Berg, E., Friedlander, M.P.: *Sparse Optimization With Least-Squares Constraints*. Department of Computer Science, University of British Columbia Technical Report TR-2010-02 (January 30, 2010) (revised January 22, 2011)
- [16] Nocedal, J., Wright, S.: *Numerical optimization*. Springer Series in Operations Research. Springer (1999)
- [17] Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn. Reson. Med.* 58, 1182–1195 (2007)
- [18] Donoho, D., Tsaig, Y., Drori, I., Starck, J.-C.: Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit. *IEEE Trans. Inform. Theory* (submitted)
- [19] Polyak, B.: *Introduction to optimization*. Optimization Software (1987) (Translated from Russian)
- [20] Tsaig, Y., Donoho, D.: Extensions of compressed sensing. *Signal Process* 86(3), 549–571 (2006)
- [21] Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory* 52, 489–509 (2006)
- [22] Candès, E., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Submitted to *Communications on Pure and Applied Mathematics* (March 2005)
- [23] Candès, E., Romberg, J.: ℓ_1 -magic: A Collection of MATLAB Routines for Solving the Convex Optimization Programs Central to Compressive Sampling (2006), <http://www.acm.caltech.edu/l1mag>
- [24] Nesterov, Y.: Gradient methods for minimizing composite objective functions. CORE Discussion Papers series, Université Catholique de Louvain, Center for Operations Research and Econometrics (2007), <http://www.uclouvain.be/en-44660.html>

Use of Imaging Techniques to Obtain 3D Models of Small Insects

Franxavier Centeno¹, Ángela López Benítez², Carles Domènech²,
Francesc Pérez-Ràfols², and Joaquim Lloveras Macià³

¹ RetoriPragmatic, P.O. Box 1476,
08755 Castellbisbal, Spain

franxavier.centeno@estudiant.upc.edu

² CDEI, C. Llorens Artigas, 4. Parc Tecnològic de Barcelona
08028 Barcelona, Spain

{lopez, domenech, f.perez}@cdei.upc.edu

³ Projectes d'Enginyeria, Edifici H (ETSEIB), planta 10. Avda. Diagonal 647
08028 Barcelona, Spain
j.lloveras@upc.edu

Abstract. This paper explores a new way to apply knowledge in the area of physics to the artistic technique. Specifically, a new application for imaging techniques in the micrometric scale is proposed: the creation of 3D digital models of small insects which would be robust and precise enough to create physical models up to 10 meters of length while maintaining a realistic appearance. X-ray computed tomography is a technique widely used in the medical world to obtain 3D images of the interior of the body non-intrusively. In this study, micro-computed tomography (microCT) is used to gain a 3D model of a 2.5mm fruit fly. From this image, reverse engineering software is used to create a polygonal model of the sample, which is then repaired until fit for reproduction in a sculpture.

Keywords: Three dimensional imaging, micro-computed tomography, small insects, additive manufacturing, art/science collaboration.

1 Introduction

A great part of insect species is still hidden from us – if not the entomology community, certainly the general population. Even for many species large enough for the human naked eye to see, their morphology is still too small to appreciate. This is the case, for example, of the fruit fly (*Drosophila Melanogaster*). This species reaches 2.5mm in length as an adult, and as such, is easily visible. However, the general morphology of its joints, mouth apparatus, compound eyes, etc. is impossible to distinguish without magnification. The objective of this paper is to propose a way to bring the insects in the boundaries of micrometric and milimetric scale closer to the general population through the use of scientific techniques and artistic practice for the construction of large sculptures.

In the past, 3D models of insects have been made by digital artists, who build the mesh from scratch with 3D modeling software such as 3DS Max, Maya or Blender. Although these meshes can be very faithful to the real insect, they are still personal interpretations from photos and 2D scans such as SEM. One of the objectives of this project is to minimize human intervention and subjectivity in the process of creating these models. For this, we propose to use a 3D imaging technique that reproduces the specimen scanned in a 3D model directly.

We have found precedents of the use of this technique to scan larger insects or other animals and create physical models using additive manufacturing [1]. However, we have not found any use of this technology for creating physical models of smaller insects, nor any center that offers physical models created for sale or renting. Therefore, we believe this new application could blur the science-art boundary, translating technical knowledge into artistic technique.

2 Imaging Technique

There are several techniques that allow the visualization of details beyond the capability of the human eye. The ideal technique for this project would be able to image details under a few micrometers, yet have a field of view of more than 3mm, so to be able to scan the whole insect without dissection. Furthermore, this technique would allow a simple reconstruction of the images into a three-dimensional digital model. For this last reason, microscopy options (optical, electron and probe microscopy) were discarded, as they only provide bidimensional images and the process to transform them into a 3D model is too complex.

Computerized tomography, on the other hand, creates volumetric models, which are relatively simple to transform into both surface and volume polygonal models. Tomography is an imaging technique based on the acquisition of unidimensional projections of the subject, which are then mathematically reconstructed into bidimensional slices. A stack of those slices is then a three-dimensional matrix representing the sample, which can be visualized or processed to obtain a polygonal model.

There are several tomographic techniques which use different physical principles to obtain the projections: X-ray (computerized tomography, CT), positrons (positron emission tomography, PET), photons (single photon emission computerized tomography, SPECT) and visible light (optical projection tomography, OPT) are some of them. Neither of the emission techniques is convenient, as they focus on functional images rather than morphological [2]. Optical tomography requires transparent samples, and although there had been studies which overcame this obstacle, the preparation is more complex than for X-ray tomography [3].

Another technique similar to tomography is magnetic resonance imaging (MRI). MRI creates slices like tomography, although the method to obtain them is not the described for tomography. MRI allows great contrast in soft tissue, but is more expensive than CT and has a slightly worse resolution [3], so finally, CT was chosen as the most appropriate technique.

CT equipment consists in an X-ray source and CCD camera coupled to a scintillator. In microCT systems such as the one used, the source emits an X-ray beam in a cone shape that illuminates the whole sample. The X-rays are attenuated by the sample and then detected by the camera. The image formed, a digital version of the conventional X-rays, is called a projection of the object. The object is then rotated and another projected image is taken. This process is repeated until there are enough projections to reconstruct the sample.

In this study, two scans by different centers were used: the first one by the *Centro di Cristallografia Strutturale* of the *Università degli Studi di Firenze*, using a Skyscan 1172 system, and the second by the *High-Resolution X-ray Computed Tomography Facility* of the *University of Texas* at Austin, using an XRadia system. Each facility also performed the reconstruction of the tomographies with the software provided with the respective scanning equipment (Nrecon and XRadia Reconstructor). The configurable parameters were chosen to emphasize the surface of the sample, foregoing contrast in its interior. They are detailed in table 1.

Table 1. Parameters used in each scan

Parameter	Firenze - Skyscan	Texas - XRadia
Objective		4X
Source voltage	59kV	50kV
Source current	167 μ A	100 μ A
Acquisition time	590ms	3s
Camera binning	1 pixels	2 pixel
Rotation	180°	190°
Views	1200	1901
Filter	No	No
Beam-hardening correction	20%	No
Reconstruction smoothing	No	Kernel size = 0.8
Final image file	8bit BMP	16bit TIFF

3 Sample Preparation

The scanning of the samples was done *in vitro*. Since the objective of the scan was a good contrast and resolution in the surface of the sample, preparation for Scanning Electron Microscopy was used. This type of fixation and staining has been extensively tested for microCT [5, 6], and although other procedures might render better contrast in soft tissue [6], that was not the target of the study. The preparation consisted on fixation with glutaraldehyde and staining with osmium tetroxide, followed by critical point drying. The detailed steps of the preparation were:

- a) Sedate the flies with chloroform for 5 minutes at ambient temperature
- b) Cleaning with ultrasound bath
- c) Fixate the flies with phosphate tampon (0.1M, pH 7.4) with 2.5% of glutaraldehyde for 60 minutes at ambient temperature and overnight at 4°C.

- d) Clean three times with phosphate tampon for 10 minutes at 4°C.
- e) Stain with osmium tetroxide 1% and ferrocyanide 0.8% in phosphate tampon for 60 minutes at 4°C.
- f) Clean three times with distilled water for 15 minutes at 4°C
- g) Dehydrate at 4°C
 - a. Ethanol 50% for 10 minutes
 - b. Ethanol 70% overnight
 - c. Ethanol 80% for 10 minutes
 - d. 3 ethanol 90% baths of 10 minutes each
 - e. 3 ethanol 96% baths of 10 minutes each
 - f. 3 ethanol 100% baths of 10 minutes each
- h) Critical point drying

4 Image Processing

Once the scan is complete, the image has to be processed to obtain the polygonal model. Different commercial software was studied to choose the one most adequate to complete each of the four steps in the acquisition of the model:

- a) Filtration of the image to improve quality
- b) 3D visualization (direct or indirect) to position the different features of the sample
- c) Segmentation of the surfaces
- d) Exportation of the 3D model

CT images might have a significant amount of noise, so filtration of these images could result in higher quality images which would make the segmentation easier. It is important that the algorithm is applied to the three-dimensional matrix, rather than the individual slices, since some structures which are clearly so when taking into account the three dimensions might appear to be noise in some individual slices.

After filtering, the volumetric model is visualized. This is done to prepare the way for segmentation, as segmenting the model slice by slice is rather time-consuming and sometimes it is difficult to differentiate features in the slices. With a good enough visualization, segmentation can begin. The software studied had slightly different methods for obtaining regions of interest and surfaces. Most offered both automatic segmentation by value and the direct drawing (with mouse) of contours, as well as region-growing algorithms.

Finally, the surface or region of interest has to be exported in a suitable format for finite element analysis or construction. It is convenient that the software used has mesh-repairing algorithms, as the conditions of the mesh for finite element analysis or additive manufacturing are very demanding.

Furthermore, to attempt the construction of the model using the method defined in section 6, the meshing algorithm used should be able to create meshes with the very specific parameters described. This algorithm would have to be developed, as the existing meshing algorithms do not match these conditions.

5 Results

The images resulting from the first scan with the Skyscan microCT equipment was a set of 793 slices, each of 1540x1540 pixels sized 2.92 μm . While the pixel size is small, the images do not have a great quality. Furthermore, as visible in the 3D surface visualization of the model in figure 1, the sample was damaged: one eye was crushed, the legs were not visible and several smaller errors diminished the overall quality of the 3D model. For this reason, it was decided to take a second scan of a different sample. The second set of images, from the Xradia equipment was a set of 382 slices, each of 992x1013 pixels. The voxels of the model were isometric, with an edge length of 4.41 μm . While the resolution of the scan (taken as the pixel size) is worse in the second scan than the first, the quality of the image is greater, as appreciated in figure 2.

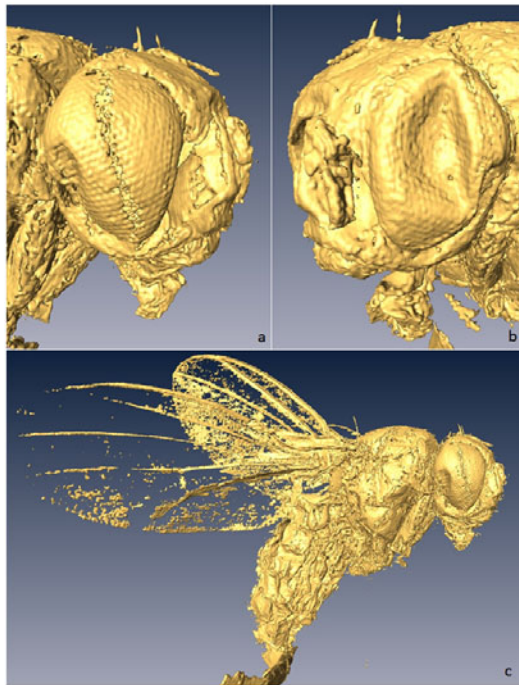


Fig. 1. Surface model of the first scan. Notice the damage left eye and the abdomen, as well as the missing limbs.

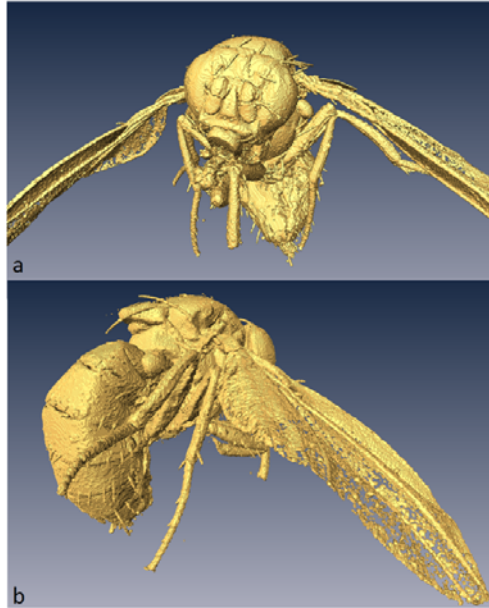


Fig. 2. Surface visualization of the second scan

6 Future Work: Physical Model

Although the first idea during the planning of this project was to build the models using additive manufacturing, it was soon discarded due to the high prices and size limitations. It was then decided to design a new manufacturing method that would allow the fabrication of different models of similar features. The method devised draws on the *trencadís*, a mosaic system proposed by the architect Antoni Gaudí, best-known representative of Catalan Modernism [7]. As shown in figure 3, Gaudí created not only patterns and images, but also 3D objects from pieces of broken tile. In this paper, we wish to propose a method to build models of this kind where fidelity in the representation of reality takes precedence over aesthetics.

The method devised must allow making sculptures of other models in the future, since the objective of this study is not the creation of one model but the design of the methodology to create several. Therefore, the materials used should be as inexpensive as possible and their availability in the future should be ensured. Ideally the material would be a waste product easy to obtain in good condition and that could be injected into a mold. One build the sculpture must withstand its own weight and the principal atmospheric phenomena: wind, rain, humidity and sunlight.

6.1 Model Structure

The general structure of the construction must be defined first. We wish to use flat pieces to build the surface of the model. If this structure was used without any



Fig. 3. Example of a sculpture by Antoni Gaudí in Parc Güell, using the *trencadís* technique. Image taken by Baikonor and hosted in Wikimedia Commons.

support, the final model would be very fragile, and possibly unable to stand its own weight. To solve this, the model will be build with a double layer in a sandwich structure. Therefore, besides the surface of the model, there will be an inner surface that will follow the shape of the outer layer at an approximately constant distance.

Hollow rigid bars will be used to attach both layers, thus giving the structure resistance. To further improve the mechanical attributes of the structure, a low-density resin will be used to fill the space between layers. A diagram of an example structure is shown in figure 4. With this kind of sandwich structure, most weight is in the outer layers (both shells), while the interior is relatively light although it occupies most of the space. Since the greatest strain the model has to stand is its own weight, this structure is fitting, as it offers good resistance while keeping its weight low.

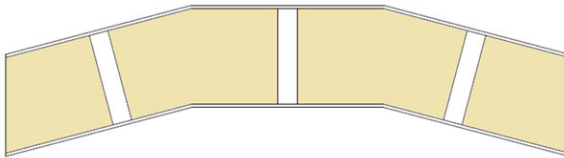


Fig. 4. Double-layer structure filled with resin foam

Initially, a topological classification as presented by Gomberg [8] was considered for selecting the geometry of the flat pieces. However, the number of necessary pieces to create a detailed model was too high, making both the creation of the pieces and their assembly too complicated. An alternative was designed using two kinds of pieces: flat pieces, which would form the surface of the model; and smaller triangular prisms, which would join two flat pieces together.

To balance ease of assembly (small number of different pieces) and resolution (larger number of pieces), it was decided best to use flat pieces in square and

triangular shapes. Square pieces would all be of equal size, of side length L , and would form the flattest sections of the surface. When the curvature of the model was greater, triangular pieces would be used. There would be a small variety of triangular pieces, but to guarantee that all pieces fit together, the length of the sides will always be multiple of L .

In order to follow the shape of the surface, in addition of using pieces of different shapes and sizes, it must be possible to unite the pieces with different relative angles between them. This is accomplished with the prism piece already mentioned. This joint consists in a prism with tabs in two of its lateral sides that would be inserted into matching slots in the borders of flat pieces. There would be joints with various angles between the two tabbed sides, thus allowing flat pieces to be joined at different angles. In figure 5, an example union between two square pieces at a near 180° angle is shown.

It is necessary to consider that the angle between the tabbed sides of the prism must be supplementary to the angle between the flat pieces. For this reason, besides the prisms with triangular bases, a prism with rectangular bases is needed to attach the pieces at the same angle. Another point to consider, especially about the unions, is their resistance to stress and their flexibility. Joints must resist tensile and compressive strain. Torque, bending and shear are expected to be relatively low. Furthermore, to facilitate assembly, the union between pieces must maintain certain level of flexibility. The solidification of the resin foam will give the final structure the final stiffness that the prism joints do not.

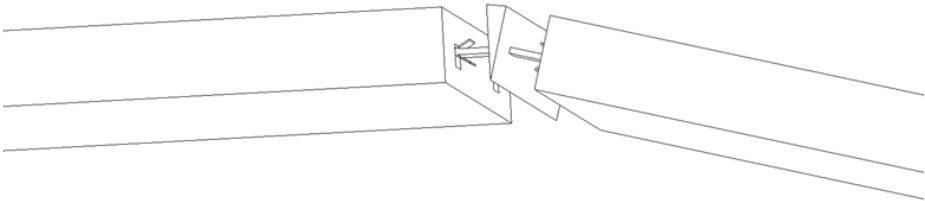


Fig. 5. Union between two flat square pieces, using a joint piece

6.2 Assembly

The structure will be assembled in a nailboard. First the pieces of the interior shell will be attached, supported by nails at the adequate height when needed. At the same time, the flat pieces will be fasted to the joints in order to shape the shell. Then, the bars will be placed. Although all pieces will have holes to place the bars (it is necessary to minimize the number of different pieces), only a concrete number of bars is necessary to maintain the structure, since the joints are responsible for keeping the shape of the surface.

Once the bars are set, the second layer of pieces will be built, using these bars as starting points (figure 6). This process will be followed, alternatively placing pieces from the interior and exterior shells as needed until closing the structure. Once the assembly is complete, valves in the bars will be used to introduce expansive foam between the shells.

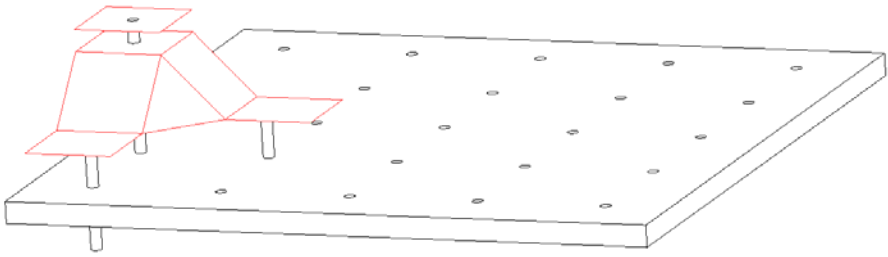


Fig. 6. Assembly of the shells of an example structure

In the case that the model is too complex, or has parts easily separated, each one could be built apart, and finally attach all sections and use the interior resin to unite the model.

7 Conclusions

At this point in the research, we find that the construction of a physical model using only imaging techniques and established manufacturing methods is more complex than expected at the beginning of the project. While additive manufacturing had to be discarded, other manufacturing methods were devised. This new proposed technique could be independent of the complexity of the sample's morphology, modeling large-scale sculptures of any organism within micrometric spatial resolutions.

We believe this project could contribute significantly to the expansion of the boundary between art and science, not only applying biomedical imaging techniques for the creation of digital models, but introducing a new technique for the construction of large-scale structures using said digital models as input.

Acknowledgments. The authors wish to thank Alfredo Balmaceda and Francisco Tejada (www.zicla.com) for their input and cooperation.

References

1. Hörnschemeyer, T., Beutel, R.G., Pasop, F.: Head Structures of *Priacma serrata* Leconte (Coleoptera, Archostemata) Inferred From X-ray Tomography. *Journal of Morphology* 252, 298–314 (2002)
2. Wernick, M.N., Aarsvold, J.N.: Emission tomography: the fundamentals of PET and SPECT. In: *Introduction to Emission Tomography*. Academic Press (2004)
3. Arridge, S.R.: *Optical Tomography in Medical Imaging* (1999)
4. Webb, A., Kagadis, G.C.: *Introduction to Biomedical Imaging*. *Medical Physics* 30, 2267 (2003)
5. Ribí, W., et al.: Imaging Honey Bee Brain Anatomy with Micro-X-Ray-Computed Tomography. *Journal of Neuroscience Methods* 171(1), 93–97 (2008)

6. Metscher, B.D.: MicroCT for Comparative Morphology: Simple Staining Methods Allow High-Contrast 3D Imaging of Diverse Non-Mineralized Animal Tissues. *BMC Physiology* 9(11), pages 11 (2009)
7. Sánchez, R.: Colour, Shine and Shape in Architectural Cladding. *Tile Today* 18(68) (2010)
8. Gomberg, B.R., et al.: Topological Analysis of Trabecular Bone MR Images. *IEEE Transactions on Medical Imaging* 19(3), 166–174 (2000)

Affine Object Tracking Using Kernel-Based Region Covariance Descriptors

Bo Ma, Yuwei Wu, and Fengyan Sun

Beijing Laboratory of Intelligent Information Technology, School of Computer Science,
Beijing Institute of Technology, Beijing 100081, P.R. China
{bma000, wuyuweii}@bit.edu.cn, sunfengyanli@163.com

Abstract. Visual tracking remains a challenging problem because of intrinsic appearance variability of object and extrinsic disturbance. Many algorithms have been recently proposed to capture the varying appearance of targets. Most existing tracking methods, however, fail to estimate the scale and orientation of the target. To deal with this problem, we model the second-order statistics of image regions using a kernel function and perform covariance matching under the Log-Euclidean Riemannian metric. Applying kernel-based covariance matrix as image region descriptor, we construct a region similarity measure that describes the relationship between the candidate object region and a given appearance template. After that, tracking is implemented by minimizing this similarity measure, in which gradient descent method is utilized to iteratively search the best matched object region. The corresponding optimization problem can be derived by calculating the first derivative of the similarity measure with respect to the affine transformation parameters and setting them to be zero. Experimental results compared with several methods demonstrate the robust performance of the proposed algorithm under challenging conditions.

Keywords: Affine object tracking, Region covariance descriptors, Log-Euclidean Riemannian metric, Gradient descent.

1 Introduction

Visual tracking is an important subject in computer vision community with a wide range of applications including visual surveillance, activity analysis, human computer interaction, and autonomous vehicles, *etc.* The goal of visual tracking is to locate a region in each frame that matches an appearance of a target. Though object tracking has been a subject of intense effort over the last two decades, it remains challenges for designing a robust tracking algorithm. Generally target representation, similarity measure and estimating the scale and orientation changes of the target are essential components of the most trackers. The contributions of this paper deal with these three components.

Recently, a new feature known as the region covariance descriptor (RCD) has been proven to be very effective tools for visual tracking, pedestrian detection, face recognition and texture classification, *etc.* By taking into account the multiple features fusion, the RCD captures both spatial and statistical properties of each pixel in an object

region using a low dimensional representation. While working effectively, the RCD is a measure of linear dependence between random variables, which may not characterize all of the structure in a given data set.

This paper presents a novel kernel-based region covariance descriptor and focuses on the discussions on its application in the affine object tracking. A kernel function is incorporated to the covariance matrix to effectively control the correlations among extracted features inside an object region in our framework. To further reduce the computation, the Log-Euclidean metric is employed as a similarity measure between covariance matrices. Based on this metric, tracking is implemented by minimizing this similarity measure, in which gradient based optimization method is utilized to iteratively search the best matched object region. At the same time, the proposed method tries to estimate the scale and orientation of the target. According to the theory of singular value decomposition (SVD), the affine matrix can be decomposed into product of two rotation matrixes and one diagonal matrix. The affine object tracking algorithm can be derived by calculating the first derivative of the similarity measure with respect to (w.r.t.) affine transformation parameters and setting them to be zero.

2 Related Work

There are three aspects of visual tracking that associate with the work presented in this paper: target representations, similarity measures, and scale and orientation changes of the target estimation. We briefly review the relevant literatures in each area.

Object representations model the characteristics of the object being tracked using features such as color, texture, edge, motion, *etc.* The literature on object representations is vast. Some of the approaches include histogram based methods [4], sparse representation based methods [11], multiple cues based methods [8] and subspace methods [13]. It is worth mentioning that the covariance region descriptor has been successfully applied to visual tracking [12]. In this framework, the tracking problem is formulated as finding a region whose covariance matrix has the smallest distance to the model covariance matrix. Porikli *et al.* [12] performed an exhaustive search in the image by comparing the given covariance template with the covariance matrix at each location using an affine-invariant Riemannian metric. Arif *et al.* [1] proposed a nonlinear covariance region descriptor for target tracking.

After achieving the target representation, the following task for tracking is to match the candidate object region and a given appearance model over consecutive frames. Many previous methods simply adopt a predefined metric including the Bhattacharyya coefficient [4], the Matusita metric [6], the affine-invariant Riemannian metric [12] and the Kullback-Leibler divergence [5]. In practice, however, predefined metric usually lacks the ability to adapt to the dynamic scene changes, since an object may change its appearance due to the pose variations, cluttered background and camera moving. Jiang *et al.* [7] described a discriminative and adaptive distance metric learning method that learns a Mahalanobis distance metric to maximize the performance of the k-NN classifier in the projected feature space. Similarly, Ma *et al.* [10] represented the distance metric between two distributions as a quadratic form, and incorporated a metric learning algorithm into the variational level set framework for accurate contour tracking.

Another challenging issue in visual tracking is how to estimate the scale and orientation of target, especially when the object is undergoing transformations. Comaniciu *et al.* [4] estimated the object scale by performing a brute-force search over a number of predefined scales and selecting the one which provides better localization. However, this method is not always suitable, and easily makes the tracker lose the target due to frequent changes in object size. Collins [3] introduced a scale-space to track changes in the object scale based on local maxima of differential scale-space filters. However, it requires high computational costs and cannot handle the rotation changes of the target. Zivkovic and Krose [17] regarded the mean shift as an EM-like algorithm, thereby presented EM-shift algorithm which simultaneously estimates the position of the local mode and the covariance matrix that can approximately describe the shape of the local mode. Zhang *et al.* [16] developed a scheme to estimate the scale and orientation changes of the object by using spatial-color features. Yilmaz [15] introduced a new kernel-based tracking using asymmetric kernels with adaptive scale and orientation selection method.

3 Kernel-Based Covariance Descriptor

Despite the covariance descriptor introduced by Porikli [12] shows good performance, in many situations the existing algorithms do not have satisfactory tracking robustness. For example, when the target undergoes scale and orientation changes, the structure distribution of the target model changes correspondingly. This is prone to losing some discriminative information and to fusing more background interference. Notice that the irregular target is localized within a rectangular window, and thus it inevitably contains some background pixels in its appearance representation. In this work, we employ an affine transformation to correctly estimate the rotation and scale of the target, which leads to minimizing the impact of background clutters.

Given the extracted feature vector $f(x)$ defined on an image plane Ω , we suppose that the image domain is divided into two parts by a curve C (e.g. rectangle, ellipse). Denote the interior region by R , the exterior region by R^c and $\Omega = R \cup R^c$, respectively. The $d \times d$ kernel covariance matrix C_R for an image region R is computed by

$$C_R(x_c) = \frac{\sum_{i=1}^N K\left(\left\|\frac{x_c - x_i}{h}\right\|\right)(f(x_i) - \mu_R)(f(x_i) - \mu_R)^T}{\sum_{i=1}^N K\left(\left\|\frac{x_c - x_i}{h}\right\|\right)}, \quad (1)$$

where μ_R is the mean vector within region R , $x_i = (x_i, y_i)_{i=1,2,\dots,N}$ are pixels in the given image region centered at $x_c = (x_c, y_c)$, h is bandwidth, respectively. And K is a kernel function with a localization property that $K(X)$ decreases and approaches zero as $|X|$ increases. In this paper, we choose the Gaussian kernel function $K(x - y) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\|x - y\|^2\right)$ as a normal kernel.

The kernel function $K_\sigma(x - y)$ is incorporated to the covariance matrix to effectively control the correlations among extracted features inside an object region in our framework. It amounts to assigning a weight for each pixel inside candidate tracking region. $K_\sigma(x - y)$ takes larger values at the point y near the center point x , and it decreases to 0 as y goes away from x . The kernel weights indicate different

contribution of each pixel in the target region, which pixels father away from the center are less reliable due to possible occlusions, interference with background, or similar boundary impacts.

Now consider the target region undergoing an affine transformation which combines shearing, scaling, rotation and translation sequentially. In our method, the geometric transformation applied on the normalized kernel is affine. In an affine transformation, the relation between the original location of a point $X = (x_i, y_i)$ and its corresponding new location $X' = (x'_i, y'_i)$ can be expressed as $X'^T = MX^T + t^T$. M is decomposed via singular value decomposition (SVD) as $M = UDV^T$, where U and V are unitary matrices which can be expressed as rotation matrices, i.e. $M = R(\phi)D(\lambda_x, \lambda_y)R(\theta)$, where $1/\lambda_x$ and $1/\lambda_y$ are positive singular values of M .

After extracting the feature vector $f(x) = f(x_i)_{i=1,2,\dots,N}$ of each pixel, the affine transformation $A = (t_x, t_y, \theta, \lambda_x, \lambda_y, \phi)$ is added into kernel-based covariance descriptor develop in Eq.(1), and the candidate image region descriptor is derived again as

$$C_R(\psi) = \frac{\sum_{i=1}^N K(\psi(A))(f(x_i) - \mu_R)(f(x_i) - \mu_R)^T}{\sum_{i=1}^N K(\psi(A))}, \tag{2}$$

where

$$\begin{aligned} \psi(A) &= \psi(t_x, t_y, \theta, \lambda_x, \lambda_y, \phi) \\ &= \frac{\left(x_c - (a(x'_i - t_x) + b(y'_i - t_y))\right)^2}{h_x^2} + \frac{\left(y_c - (c(x'_i - t_x) + d(y'_i - t_y))\right)^2}{h_y^2}. \end{aligned} \tag{3}$$

4 Affine Object Tracking

4.1 Region Similarity Measure

Since a nonsingular covariance matrix is a symmetric positive definite (SPD) matrix lying on a connected Riemannian manifold, statistics for covariance matrices may be computed through Riemannian geometry. We adopt the Log-Euclidean Riemannian metric in this paper. Under this metric, the distance between two points X and Y is calculated by $\|\log(Y) - \log(X)\|$. Thus, the region similarity measure can be defined as

$$\rho(C_R(\psi(A)), C_T) = \|\log C_R(\psi) - \log C_T\|_F, \tag{4}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix.

Given an image, we aim to find a target region centered at (x_c, y_c) and to estimate the scale and orientation of the target in the image sequence by minimizing the similarity measure between the candidate object region C_R and a given appearance model C_T given in Eq.(4). The initialization of tracking is set by manually marking the target object box in the first frame. Once this box is localized, the resultant kernel

covariance image descriptor is saved as an appearance model C_T . After that, gradient descent method is utilized to iteratively search the best matched target region for every subsequent frame.

4.2 The Optimization Procedure

To minimize the Eq.(4), we need to compute its first variation w.r.t. $t_x, t_y, \theta, \lambda_x, \lambda_y$ and ϕ . Let $\rho(C_R(\psi), C_T) = \left[\sum_{i=1}^d \sum_{j=1}^d \rho^2(C_R(\psi), C_T) \right]^{1/2}$, where d represents the dimensional of feature vector. For the derivative of $\rho(C_R(\psi(A)), C_T)$ w.r.t. affine transformation parameters $A = (t_x, t_y, \theta, \lambda_x, \lambda_y, \phi)$, we have

$$\frac{\partial \rho}{\partial A} = \frac{1}{\rho} \sum_{i=1}^d \sum_{j=1}^d \rho_{i,j} \left(\frac{\partial \rho_{i,j}}{\partial A} \right)_{i,j} = \frac{1}{\rho} \sum_{i=1}^d \sum_{j=1}^d \rho_{i,j} \left(C_R^{-1}(\psi(A)) \frac{\partial C_R(\psi(A))}{\partial A} \right)_{i,j}. \quad (5)$$

Now, we need to compute $\frac{\partial C_R(\psi(A))}{\partial A}$. Let $K(\psi(A)) = K(\psi(t_x, t_y, \theta, \lambda_x, \lambda_y, \phi))$. We obtain

$$\begin{aligned} \frac{\partial C_R(\psi)}{\partial A} &= \frac{\sum_{i=1}^N \left(K'(\psi(A)) \cdot (f(x_i) - \mu_R)(f(x_i) - \mu_R)^T \right) \cdot \sum_{i=1}^N K(\psi(A))}{\left(\sum_{i=1}^N K(\psi(A)) \right)^2} \\ &\quad - \frac{\sum_{i=1}^N \left(K(\psi(A)) \cdot (f(x_i) - \mu_R)(f(x_i) - \mu_R)^T \right) \cdot \left(\sum_{i=1}^N K(\psi(A)) \right)'}{\left(\sum_{i=1}^N K(\psi(A)) \right)^2}, \end{aligned} \quad (6)$$

where $K'(\psi(A))$ denotes the first derivative of Gaussian kernel function $K(\psi(A)) = (2\pi)^{-d/2} \exp(-\frac{1}{2}\|\psi(A)\|^2)$ w.r.t. affine transformation parameters A .

One can respectively compute the first derivative of $C_R(\psi(A))$ w.r.t. affine transformation parameters. Substituting $\frac{\partial C_R(\psi(A))}{\partial A}$ into Eq.(5) leads to the gradient descent equation of Eq.(4). That is, the iterative formulation for calculating the optimal parameter of A is given by

$$(A)_{t+1} = (A)_t - \frac{\Delta t}{\rho} \sum_{i=1}^d \sum_{j=1}^d \rho_{i,j} \left(C_R^{-1}(\psi(A))_t \frac{\partial C_R(\psi(A))_t}{\partial A} \right)_{i,j}. \quad (7)$$

This equation provides an iterative procedure to seeking the optimal affine transformation parameters A .

4.3 Template Updating

The object is tracked by manually marking a template in the first frame and finding the region of interest in consecutive frames. Since the foreground targets undergo shape, size, and appearance variations, it is difficult for a fixed appearance template to

adapt to these variations. We update the template by computing the Log-Euclidean Riemannian mean of covariance matrices. The computation is simply Euclidean in the logarithmic domain, and the final mean is computed by mapping back to the Riemannian domain with the exponential map.

Specifically, we use the Log-Euclidean mean of n covariance matrices with positive weights $(w_i)_{i=1}^n$ such that $\sum_{i=1}^n w_i = 1$ is a direct generalization of the geometric mean of the matrices. The weight of each frame is dependent on the similarity with the previous frame. It is computed by $\chi = \frac{1}{d} \|\log C_i - \log C_n\|^{-1}$ where $d = \sum_{i=1}^n \|\log C_i - \log C_n\|^{-1}$, C_n denotes the current covariance matrix. During the template updating, we keep a set of n previously obtained covariance matrices C_1, C_2, \dots, C_T . From this set, we compute a sample mean covariance matrix as a new template. The weighted covariance mean algorithm can be summarized in Algorithm 1.

Algorithm 1. Weighted covariance mean algorithm based on Log-Euclidean metric

Input:

- The iteration error threshold ε ;
- The covariance matrices $C_i (i = 1, 2, \dots, n)$;

Output:

Mean μ ;

1. Initialize the mean $\mu = C_1$;
 2. Compute $d = \sum_{i=1}^n \|\log C_i - \log C_n\|^{-1}$;
 3. **while** $\|\log(\Delta\mu)\| \geq \varepsilon$
do
 - $\Delta\mu_i = \mu^{-1} C_i$
 - $\Delta\mu = \exp\left(\frac{1}{d} \|\log C_i - \log C_n\|^{-1} \log(\Delta C_i)\right)$
 - $\mu = \mu \cdot \Delta\mu$**end**
 4. **return** mean μ
-

5 Experimental Results

To evaluate the performance of the proposed algorithm, we run our proposed tracking algorithm on several challenging video sequences, where the targets undergo occlusions, lighting changes, scale variations, object rotations and complex backgrounds, *etc.* For comparisons, we implement five state-of-the-art tracking algorithms including IVT [13], generalized mean shift visual tracking [14] (referred to as GMS tracker), GPF [9], MIL [2], and the covariance tracking [12] (referred to here as Cov tracker). We present some representative frames to show the performances of the proposed tracker and the other trackers.

Rotation and partial occlusion: The first test sequence shows that a girl's head is continuously moving and rotating from left to right, resulting in large variation in size and pose angle, together with an partial occlusion. From the tracking results in Fig.1, we can observe that MIL tracker fails after the intersection at around the frame 98 due to its inability to handle the scale and the orientations of. It can be explained that the MIL tracker mistakenly updates its model with most recent instances from image patches of the girl's face. The other trackers manage to track the object till the end of the video. However, at some frames (such as around the frames 120, 190 and 226), tracking accuracy is noticeably worse than those of our tracker.



Fig. 1. Tracking results of the *Girl 2* sequence. From top to bottom are the results of IVT [13], MIL [2], GPF [9], GMS [14], Cov [12] and our method, respectively.

Pose variations: The *Female skater* sequence contains over 150 frames, and the dazzling performance is accompanied by an large pose variation. Fig.2 demonstrates some representative tracking results. While the background is relatively simple, IVT and Cov trackers are not able to locate the skater accurately. When the tracked skater turns her body the IVT tracking box deviates from the target center at the frame 62 (e.g. part of the object is outside the box), and then together with the ongoing pose variation it gets lost at the frame 90. Though GMS and GPF trackers track the whole sequence successfully, the tracking accuracy is quite low, especially when the body turn around (e.g. the frames 122 and 158). Overall, our method is able to track the skater well and provides tracking boxes that are much more accurate and consistent. The MIL tracker also obtains good results.



Fig. 2. Tracking results of the *Female skater* sequence. From top to bottom are the results of IVT [13], MIL [2], GPF [9], GMS [14], Cov [12] and our method, respectively.

Illumination and scale changes: several skaters perform on stage with drastic lighting change as a result of neon and spot lights, as shown in Fig.3. During tracking, the object template image has been normalized to have unit variance and zero mean so as to minimize the effects of illumination changes. However, the most methods (e.g. IVT tracker, MIL tracker and GMS tracker) cannot track the target correctly. GPF tracker fails at the frame 75; recovers at the frame 220 and then fails again. The GMS trackers employed the color histogram as object representation, leading to inability to handle the case of lighting condition drastically changed. Overall, our method outperforms the other trackers. Nevertheless, our tracker eventually fails after the frame 234 as a result of a combination of drastic pose and illumination change. After the frame 234, tracking such a target is extremely difficult because the target is almost indistinguishable in the dark environment, even for human eyes.

Severe occlusion: From Fig.4, we see that the MIL tracker drifts when the long-term occlusion happens at the frame 145. The appearance model fuses more background interference due to an occlusion, which significantly influences the instances online update of the MIL tracker. GPF and Cov trackers lose the target at the frame 145 and frame 217 respectively, analogously to the MIL tracker. GMS tracker gives the results of tracking the upper body of the walking woman, using code provided by the corresponding author. In contrast, our method and IVT tracker achieve good results. The tracking error gradually becomes larger, but overall, the results remain acceptable.

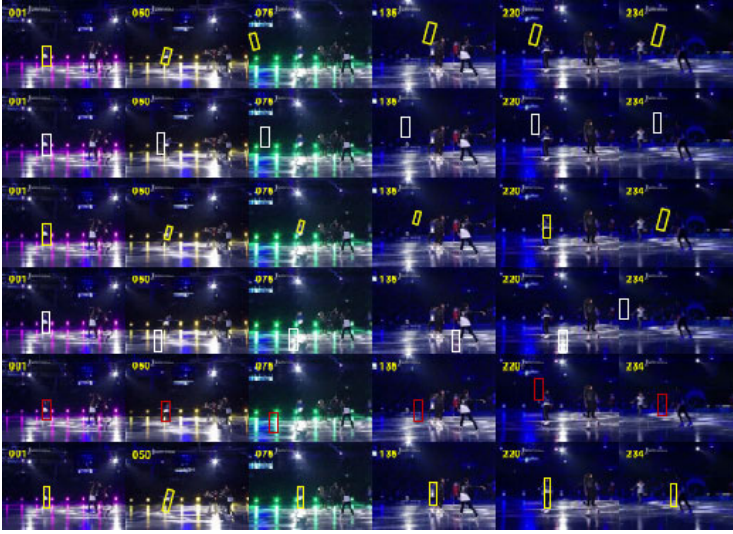


Fig. 3. Tracking results of the *skating1* sequence. From top to bottom are the results of IVT [13], MIL [2], GPF [9], GMS [14], Cov [12] and our method, respectively.



Fig. 4. Tracking results of the *Walking woman* sequence. From top to bottom are the results of IVT [13], MIL [2], GPF [9], GMS [14], Cov [12] and our method, respectively.

Quantitative comparisons: To quantitatively measure the accuracy of the tracking results under challenging conditions, we manually mark the ground truth by selecting the minimal window that encloses the target in every three frames for the tested sequences. The error metric adopted for comparison is based on the relative position

errors (in pixels) between the center of the tracking results and that of the ground truth. Ideally, an optimal tracker is expected to have a small error. For the *Female Skater* sequence, we do not include upper limb parts in the ground truth when they are spread out too much. The quantitative comparison results are displayed in Fig.5. We see that the position differences of the results in our method are smaller than those of the other trackers.

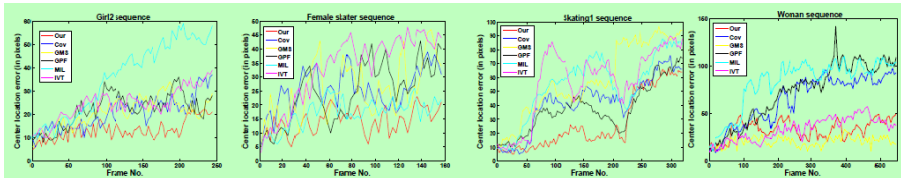


Fig. 5. Quantitative comparison of the trackers in terms of position errors (in pixels)

6 Conclusion

In this paper, we have introduced an algorithm for accurate tracking of objects undergoing pose variation, long-term occlusion, affine object transformations and illumination changes. Under the Log-Euclidean Riemannian metric, kernel-based region covariance descriptor is employed to characterize the object appearance. After that, tracking is implemented by minimizing Log-Euclidean Riemannian metric, in which gradient descent optimization method with respect to affine transformation parameters is utilized to iteratively search the best matched object region. The proposed method is capable of estimating the scale and orientation of the target. For further robustness, we dynamically update the template by computing the geometric mean of covariance matrices in Riemannian manifold. In our experiments involving several challenging sequences and five other state-of-the-art trackers, our tracking algorithm demonstrates promising performances.

References

1. Arif, O., Vela, A., et al.: Non-rigid object localization and segmentation using eigenspace representation. In: IEEE 12th International Conference on Computer Vision, pp. 803–808 (2009)
2. Babenko, B., Yang, M., Belongie, S.: Robust object tracking with online multiple instance learning. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(8), 1619–1632 (2011)
3. Collins, R.: Mean-shift blob tracking through scale space. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 234–240 (2003)
4. Comaniciu, D., Ramesh, V., Meer, P.: Kernel-based object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(5), 564–575 (2003)
5. Elgammal, A., Duraiswami, R., Davis, L.: Probabilistic tracking in joint feature-spatial spaces. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 781–788 (2003)

6. Hager, G., Dewan, M., Stewart, C.: Multiple kernel tracking with ssd. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 790–797 (2004)
7. Jiang, N., Liu, W., Wu, Y.: Learning adaptive metric for robust visual tracking. IEEE Transactions on Image Processing (in press, 2011)
8. Kuo, C., Huang, C., Nevatia, R.: Multi-target tracking by on-line learned discriminative appearance models. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 685–692 (2010)
9. Kwon, J., Lee, K.M., Park, F.: Visual tracking via geometric particle filtering on the affine group with optimal importance functions. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 991–998 (2009)
10. Ma, B., Wu, Y.: Learning distribution metric for object contour tracking. In: Proceedings of International Conference on Multimedia Technology (ICMT 2011) (to appear, 2011)
11. Mei, X., Ling, H.: Robust visual tracking using l1 minimization. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 1–8 (2009)
12. Porikli, F., Tuzel, O., Meer, P.: Covariance tracking using model update based on lie algebra. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 728–735 (2006)
13. Ross, D., Lim, J., Lin, R., Yang, M.: Incremental learning for robust visual tracking. International Journal of Computer Vision 77(1), 125–141 (2008)
14. Shen, C., Kim, J., Wang, H.: Generalized kernel-based visual tracking. IEEE Transactions on Circuits and Systems for Video Technology 20(1), 119–130 (2010)
15. Yilmaz, A.: Kernel Based Object Tracking Using Asymmetric Kernels with Adaptive Scale and Orientation Selection. Machine Vision and Applications 22(2), 255–268 (2011)
16. Zhang, H., Huang, W., Huang, Z., Li, L.: Affine object tracking with kernel-based spatial color representation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 293–300 (2005)
17. Zivkovic, Z., Krose, B.: An em-like algorithm for color-histogram-based object tracking. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 798–803 (2004)

A Pulmonary Nodules Detection Method Using 3D Template Matching

Ting Gao¹, Xiwen Sun², Yuanjun Wang¹, and Shengdong Nie¹

¹School of Medical Instrument & Food Engineering,
University of Shanghai for Science and Technology,
200093, Shanghai, China

tinagog2525@yahoo.cn, wyj803@yahoo.com.cn, nsd4647@163.com

²Radiology Department,
Shanghai Pulmonary Hospital,
200093, Shanghai, China
xwsun40@smmail.com

Abstract. A pulmonary nodules detection algorithm based on 3D template matching method using CT images is proposed. Firstly, lung parenchyma was segmented from CT series images. Secondly, apply the high-pass filter on the lung parenchyma images to enhance the edge of lung nodules. Then, design 3D templates with different size based on the nodules' features. At last, apply the 3D-SSD (sum of squared differences) template matching algorithm between the 3D templates and the lung fields, and the final matching results were labeled as lung nodules on the original images. Using 20 clinical data set (include 35 pulmonary nodules in 3-20mm) to test the detection method, the accuracy rate is 81.08%, false positive rate (FP) is 5.4% and sensitivity rate is 85.71%. Therefore, the pulmonary nodules detection algorithm proposed in this paper can detect different typological nodules accurately and effectively.

Keywords: pulmonary nodules detection, edge enhancement, 3D template matching, SSD.

1 Introduction

The rate of lung cancer mortality ranks on top position for all cancers. A main reason for the high mortality is that early diagnosis of lung cancer is very difficult while it's hardly treatable in terminal stage. However, lung cancer commonly appears pulmonary nodules in early stage. Therefore, the early clinical diagnosis generally begins with the pulmonary nodules detection using CT images. Pulmonary nodules usually present as sphere-shaped lesions in CT images, smaller than 30mm in diameter. Pulmonary nodules can be divided into two categories: solid nodule and sub-solid nodule. A solid nodule can be defined as a nodule that completely obscures the entire lung parenchyma within it. The sub-solid nodule contains part-solid nodule and non-solid nodule. Part-solid nodules are those having sections that are solid in this sense, while non-solid nodules are those with no solid part. To detect solid nodule is ease comparatively, because it has distinct boundary and higher contrast. However, due to the sub-solid

nodules usually have indistinct boundaries and lower contrast, to detect sub-solid nodule is very difficult.

Many literatures have been reported about using CAD (computer aided detection) technique to detect pulmonary nodules with CT images in the past several years. CAD algorithm study for nodules detection is a hot research topic currently in medical imaging processing field. There are two main approaches for pulmonary nodules detection: classifier-based method and template matching ones. The former one with widely used, according to some proposed methods[1] [2], yields a high accuracy rate but also producing a high false-positive(FP) rate. Besides, this method costs high complex computation[3]. The advantage of the later one is the simplicity of matching template and detection algorithm design. Meanwhile, it's faster and more efficient than methods based on classifier. But this method has the low accuracy rate as well as high FP rate. So far, all methods reported in home and abroad can only detect solid nodules.

There are some of approaches derived from template matching-based method, including those based on the deformation of the object, and methods based on anatomy models[4][5], et.al. 2001, Lee et.al. proposed a novel template matching technique based on a genetic algorithm (GA) template matching (GATM) for detecting solid nodules, it is fast with the accuracy of 72% but with high FP[6]. In 2004, Farag et.al. proposed an algorithm for nodule detection using deformable 3D and 2D templates describing typical geometry and gray level distribution within the nodules of the same type. The detection combines normalized cross-correlation (NCC) template matching by genetic optimization and Bayesian post-classification. The accuracy is 82.3% with an FP of 7.9%[7]. In 2007, Wang et.al. changed layers' model of the 3D templates proposed by Lee, achieving a higher accuracy[8]. In 2006, A. Kohandani et.al. designed a method based on gray level changes, which has a good sensitivity rate, 92.3%[9]. The next year, Onur Osman et.al. found a new way to reduce FP to 9.2% base on that method [10]. In 2007, Jamshid et.al. designed templates based on the spherical enhancement, using 3D geometric feature calculation and the shape based on GATM template matching algorithm, achieving an accuracy of some 90% with an FP of 9.8%[11]. In 2010, Joao Rodrigo et.al. make 3D reconstruction design on the basis of the template matching method, and the accuracy of that method is 84.84%[12].

Until now, most of the detection approaches based on template matching use a regular sphere shape, these templates can get a good detection result for solid nodules with a similar regular shape, while it is hard to detect the non-solid ones with anomalous shape and big edge difference. Considering this shortage, we improved the current existent template based detection method. On the one side, we classify these different types of nodules and apply to the training of the designed sphere templates, which improve the similarity between the templates and different pulmonary nodules. On the other side, we generalize the SSD matching method from 2D to 3D, which further decrease the FP rate of detection. Experimental results prove that the designed 3D templates based detection algorithm is valid for different types of pulmonary nodules.

2 Design for 3D Template Matching Algorithm

2.1 3D Pulmonary Nodules Template Matching Method

The steps of the proposed 3D template matching method for pulmonary nodules detection are described as follows: Firstly, segments lung parenchyma from CT series images effectively and accurately using the method of threshold. Secondly, apply the high-pass filter on the lung parenchyma segmented from CT series images to enhance the edge of nodules. Thirdly, design the 3D spherical templates of different size and use the features of nodules to train them. Fourthly, extend the 2D template matching method SSD to 3D and apply the 3D-SSD template matching algorithm between the 3D templates and the lung fields, and the final matching results were labeled as lung nodules on the original images.

Eight 3D templates with different sizes have been designed, ranging from 3mm to 20mm in diameter. The algorithm will do the matching one by one automatically. The flow chart shows in Fig. 1.

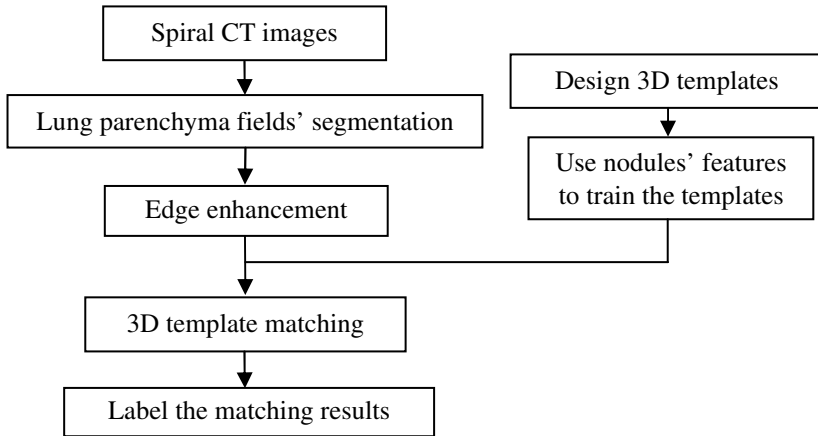


Fig. 1. Flow chart for the 3D pulmonary nodules template matching method

2.2 Lung Parenchyma Fields Segmentation

The lung parenchyma fields are isolated first to control the scope and improve the precision of the matching. The process of lung parenchyma fields are segmented according the following steps.

- (a) Use Otsu's algorithm to compute a threshold[13][14], use the threshold to get the binary image Fig. b.
- (b) Remove the bed and body part with inverse operation and some semi-automatic way, get Fig. c.
- (c) Use open and close morphological operator based on Fig. c, get the lung parenchyma mask Fig. d.
- (d) Apply the mask on the original image, get lung parenchyma fields image Fig.e.

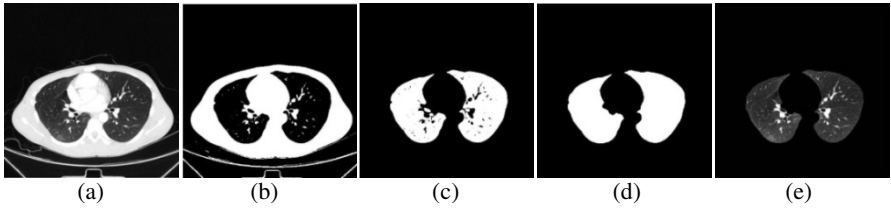


Fig. 2. Lung parenchyma fields segmentation method(a) original image (b)binary image(c)remove bed and body part(d)after morphological processing(e) lung parenchyma fields' image

2.3 Nodule Edge Enhancement

In 2007, Jamshid et.al.[11] designed a spherical characteristics enhancement method for template matching. Feature enhancement can make the edges of the sub-solid pulmonary nodules clearer, facilitating detection of them. This paper uses high-pass filter to enhance edges of the nodules. This method is simple and fast with a satisfactory result [15].

High-pass filter algorithm put filter F upon the data D , then do the template filter processing (Equation (1)). M is the width of F ; N is the height of F [16].

Here we use image as D , $F = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$.

$$Result(i, j) = \sum_{n=1}^N \sum_{m=1}^M D(i + m, j + n) * F(m, n) \tag{1}$$

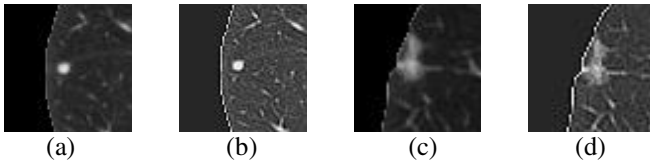


Fig. 3. Edge enhancement result (a) an example of solid nodule (b) edge enhance on a (c)an example of sub-solid pulmonary nodule (d) edge enhance on c

From above enhancement examples, we can see the high-pass filter algorithm can enhance the edge for both solid-nodule and sub-solid nodule efficiently.

2.4 3D Spherical Templates

Many different types of 3D templates have been designed in previous research. With the development of CT technology, much thinner layer can be acquired. In order to design a much similar template with the nodules, here a new template training method is employed, which utilizes the features of pulmonary nodules. Spherical templates are made in 3-20mm in diameter, with the number of layers the same as average, the count of the nodules is shown in Table 1:


























Table 1. Training nodules' details

Diameter(mm)	3	5	8	10	12	15	18	20
Count of sub-solid nodules(total :23)	1	3	5	5	4	3	1	1
Count of sub-solid pulmonary nodules(total :15)	0	1	2	3	3	4	2	0
The number of layers in average	2	5	8	10	12	16	18	21

The training process of 3D pulmonary nodules templates is according the following steps, suppose the diameter of the spherical solid nodule is 5mm.

- (a) Create an empty 3D spherical template with the same layers as the average layers of 5mm nodules. The center area of the template is a spherical of 5mm in diameter, and the background value is set according the lung parenchyma gray intensity.
- (b) Collect the features of all known 5mm nodules. The features are classified by shape and type “*solid*” or “*non-solid*”. We choose 3 of them by the same feature of “*spherical*” and “*solid*”. They are Feature 1, 2 and 3 in Table 2.
- (c) Reset the value of the center part of the template according the average of the feature, ensure the operation is within the spherical area.
- (d) Make a Gaussian blur on the edge of the spherical template.

Table 2. 3D pulmonary nodules template training example

Empty template					
Feature 1					
Feature 2					
Feature 3					
Result					

2.5 3D-SSD Template Matching Method

A template matching algorithm was used to search the object from image S . At first, we get the feature of the object T , and make the template called T . Then, we use the algorithm to calculate the similarity of the S and T , finding the most similar place, and get the location of the object.

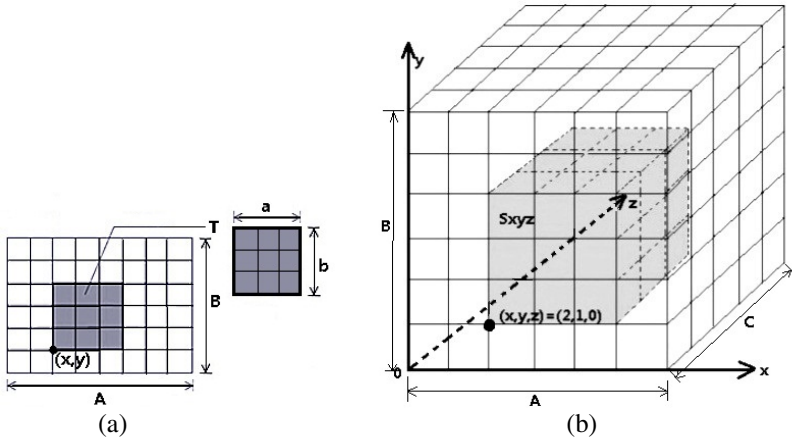


Fig. 4. The template matching algorithm (a) 2D mode (b) 3D mode

2D template matching method was widely used [17]. As can be seen in Fig.4(a), $T(a \times b)$ has been shifted inside $S(A \times B)$, T covers the area S_{xy} , as we see the colored region in the figure. Here (x, y) is the location of the most left, bottom position of S_{xy} inside S . We use planar sum of squared differences (2D-SSD, Equation (2)) to calculate the similarity of S_{xy} and T .

The template matching method is generalized to 3D mode here: $T(a \times b)$ expand to $T(a \times b \times c)$, $S(A \times B)$ expand to $S(A \times B \times C)$. At the same time, S_{xyz} is the covered area instead of S_{xy} and (x, y, z) is the location of the most left, bottom and top position of S_{xyz} inside S . 3D-SSD (Equation (3)) is used here to calculate the similarity of S_{xyz} and T . Various difference measures have different mathematical properties, and different computational properties. Here we use a fast template matching algorithm SSD to calculate and evaluate the computational properties, and expand it from 2D to 3D to complete 3D template matching. The range of the SSD value is (0,1) both of the 2D mode and the 3D mode, the larger the value is, the more similar the template is. [18]

$$2D-SSD(x, y) = \sum_{i=1}^a \sum_{j=1}^b [S(x + i, y + j) - T(i, j)]^2. \tag{2}$$

$$3D-SSD(x, y, z) = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c [S(x + i, y + j, z + k) - T(i, j, k)]^2. \tag{3}$$

The range of the 2D search is $\begin{cases} 1 \leq x \leq A - a \\ 1 \leq y \leq B - b \end{cases}$.

The range of the 3D search is $\begin{cases} 1 \leq x \leq A - a \\ 1 \leq y \leq B - b \\ 1 \leq z \leq C - c \end{cases}$.

3 Results and Evaluation

The image data used in this paper was acquired from Department of Radiology, Shanghai Pulmonary Hospital, collected by Philips 40 row helical CT with a layer thickness of 2mm, 0.84mm in pixel size. This method can detect the pulmonary nodules with 3-20mm in diameter.

Using 20 clinical data set (include 35 pulmonary nodules in 3-20mm) to test the detection method, 24 of 35 are solid nodules and the others are sub-solid ones. Data close by the vessel and lung walls are included.

The evaluation index of this pulmonary nodules' detection method includes the *Accuracy*, *Sensitivity*, *False Negatives (FN)* and *False Positive (FP)*. High level of Accuracy rate, Sensitivity rate and low FP rate represent a good detection method.

$$Accuracy = \frac{TP}{TP+FP+FN}. \quad (4)$$

$$Sensitivity = \frac{TP}{TP+FN}. \quad (5)$$

$$FN = 1 - Sensitivity. \quad (6)$$

Table 3. The result of this 3D pulmonary nodules detection template matching method

Nodules exist	Nodules detected	Result			Accuracy	Sensitivity	FP rate
		TP	FP	FN			
35	32	30	2	5	81.08%	85.71%	5.4%

As we can see from Table 3, this 3D pulmonary nodules detection template matching method has high sensitivity, accuracy and low false positive rate, reason of false positive mainly caused by the part of the wall near the lung. It achieves a better result than the existing template matching algorithms, and can provide a powerful tool for doctors.

4 Conclusion

The experimental results show that this 3D pulmonary nodules detection template matching method has high sensitivity, accuracy and low false positive rate; it can be used in doctors' nice assistance to reduce the misdiagnosis. This detection method is stable, reliable, and independent of image grey values. So it can assist doctors effectively to detect pulmonary nodules, which is beneficial to the actual clinical application of CAD of lung cancer.

Acknowledgments. This work is under the special auspice of National Natural Science Foundation (No. 60972122) and Shanghai Municipal Education Commission Research and Innovation Project (No. 09YZ216).

Ting Gao, she is currently a candidate for her master degree in University of Shanghai for Science and Technology. Her main research interest is medical image processing. (E-mail: tinagog2525@yahoo.cn).

Xiwen Sun, Ph.D., he is a professor from Tongji University and director of Radiology Department, Shanghai Pulmonary Hospital, his main research interests are lung cancer diagnosis using CT/MRI. (E-mail: xwsun40@smmail.com).

Yuanjun Wang, Ph.D., he is now a lecture in University of Shanghai for Science and Technology. His research interest is medical image processing and analysis. (E-mail: wyj803@yahoo.com.cn).

ShengdongNie, Ph.D., professor, whose research interests include medical image processing, Computer-Aided Diagnosis(CAD) based on medical image, the corresponding author of this paper, School of Medical Instrument & Food Engineering, University of Shanghai for Science and Technology, Shanghai, 200093, China. (phone:86-021-55271172; fax: 86-021-55620106; E-mail: nsd4647@163.com).

References

- [1] Lee, S.L.A., Kouzani, A.Z., Hu, E.J.: Automated detection of lung nodules in computed tomography images: a review. In: *Machine Vision and Applications*, May 14 (2010)
- [2] Messay, T., Hardie, R.C., Rogers, S.K.: A new computationally efficient CAD system for pulmonary nodule detection in CT imagery. *Medical Image Analysis* 14(3), 390–406 (2010)
- [3] Saba, L., Caddeo, G., Mallarini, G.: Computer-aided detection of pulmonary nodules in computed tomography: analysis and review of the literature. *Journal of Computer Assisted Tomography* 31(4), 611–619 (2007)
- [4] Zhang, H., Chen, Z.: Research on Classical Image Matching Algorithm and Its Improved Method. *Exploitation and Application of Software* 27(9) (2008)
- [5] Hong, H., Lee, J., Yim, Y.: Automatic lung nodule matching on sequential CT images. *Computers in Biology and Medicine* 38(5), 623–634 (2008)
- [6] Lee, Y., et al.: Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. *IEEE Transactions on Medical Imaging* 20(7), 595–604 (2001)
- [7] Farag, A.A., et al.: Detection and recognition of lung abnormalities using deformable templates. In: *Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004*, vol. 3, pp. 738–741 (2004)
- [8] Wang, P., et al.: Lung metastases detection in CT images using 3D template matching. *Medical Physics* 34(3), 915 (2007)
- [9] Ozekes, S., Camurcu, A.Y.: Automatic Lung Nodule Detection using Template Matching. In: Yakhno, T., Neuhold, E.J. (eds.) *ADVIS 2006*. LNCS, vol. 4243, pp. 247–253. Springer, Heidelberg (2006)
- [10] Osman, O., Ozekes, S., Ucan, O.N.: Lung nodule diagnosis using 3D template matching. *Computers in Biology and Medicine* 37(8), 1167–1172 (2007)
- [11] Dehmeshki, J., et al.: Automated detection of lung nodules in CT images using shape-based genetic algorithm. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society* 31(6), 408–417 (2007)
- [12] Da Silva Sousa, J.R.F., et al.: Methodology for automatic detection of lung nodules in computerized tomography images. *Computer Methods and Programs in Biomedicine* 98(1), 1–14 (2010)

- [13] Otsu, N.: A threshold selection method from gray level histogram. *IEEE Trans Systems Man Cybernetics* 8, 62–66 (1978)
- [14] Li, Z.: Simple image segmentation threshold method. *Science and Technology Innovation Herald* 11 (2010)
- [15] Zhang, C.: Image enhancement technology research based on the Matlab. *Science and Technology Information* 21 (2009)
- [16] Gonzalez, R.C., Woods, R.E., Eddins, S.L.: *Digital Image Processing use MATLAB*, pp. 99–101. Publishing House of Electronics Industry (2009)
- [17] Zuo, L.: Imaging Progressing—Template Matching Method. *Image Project*, 163–165 (2002)
- [18] Atallah, M.J.: Faster image template matching in the sum of the absolute value of differences measure. *IEEE Transactions on Image Processing* 10(4), 659–663 (2001)

An Improved Iterative Binary Coloring Procedure for Color Image Segmentation

Javier Montero¹, Susana Muñoz¹, and Daniel Gómez²

¹ Dpto. de Estadística e Investigación Operativa, Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid, 28040 Madrid, Spain

² Dpto. de Estadística e Investigación Operativa III, Escuela de Estadística, Universidad Complutense de Madrid, 28040 Madrid, Spain
javier_montero@mat.ucm.es, {smunoz,dagomez}@estad.ucm.es

Abstract. In this work we present an improvement on an iterative binary coloring procedure for image segmentation taken from the literature. We introduce some modifications in the way of dealing with the so-called inconsistent pixels, and we show the results obtained by applying both procedures to a satellite image of the province of Seville. The computational experience that we have performed shows that, in general, the modified procedure leads to images of similar or better quality than the ones obtained by the original procedure, as well as to a significant reduction of the number of final regions.

Keywords: Segmentation techniques, Graph theory, Decision support systems.

1 Introduction

The segmentation of a digital image consists in partitioning the image into several regions, in such a way that all the pixels belonging to the same region are similar in a certain way.

This problem has been widely studied in the literature, and it has been tackled by using different procedures, such as Markov random fields (see e.g. [14]), adaptive neuro-fuzzy systems (see [1]), histogram-based approaches (see e.g. [18,3]), the Expectation-Maximization algorithm (see [2]), the k -means clustering algorithm (see e.g. [16,12]), and graph-based methods (see e.g. [19,17,15,4,9,10,6,7,20,13]).

In this work we introduce some modifications in the relaxed coloring algorithm proposed in [7]. The main idea of that algorithm was also used in [6,20,8], and it basically consists in applying iteratively a binary coloring process to the pixels of the image under consideration, taking into account a prefixed threshold in each iteration.

This paper is organized as follows: In Sect. 2 we give the basic concepts and notation that we are going to use. In Sect. 3 we outline the iterative binary coloring procedure proposed in [7]. In Sect. 4 we introduce some improvements in that procedure. In Sect. 5 we show the results obtained by applying both procedures to a satellite image. In Sect. 6 we give some final comments.

2 Basic Concepts and Notation

In this section we review the basic concepts and notation used throughout the paper (see [7] for more details).

A digital image is understood as a rectangular set of pixels. Each pixel (i, j) is characterized by the values x_{ij}^1 , x_{ij}^2 and x_{ij}^3 of the red, green and blue band of the visible spectrum, respectively. The information about an image can be represented as $I = \{(x_{ij}^1, x_{ij}^2, x_{ij}^3) \mid (i, j) \in P\}$, where $P = \{(i, j) \mid i \in \{1, \dots, r\}, j \in \{1, \dots, s\}\}$ is the set of pixels and $r, s \in \mathbb{N}$ (we say that the size of this image is $r \times s$).

We consider the undirected graph $G(I) = (P, E_1 \cup E_2)$, where E_1 and E_2 are the sets of all possible vertical and horizontal edges, respectively, i.e.,

$$E_1 = \{(i, j), (i + 1, j)\} \mid i \in \{1, \dots, r - 1\}, j \in \{1, \dots, s\}$$

and

$$E_2 = \{(i, j), (i, j + 1)\} \mid i \in \{1, \dots, r\}, j \in \{1, \dots, s - 1\}.$$

We associate the Euclidean distance $d_{ij}^1 = \sqrt{\sum_{k=1}^3 (x_{ij}^k - x_{i+1,j}^k)^2}$ to each edge

$\{(i, j), (i + 1, j)\} \in E_1$, and the Euclidean distance $d_{ij}^2 = \sqrt{\sum_{k=1}^3 (x_{ij}^k - x_{i,j+1}^k)^2}$

to each edge $\{(i, j), (i, j + 1)\} \in E_2$, and we consider the resulting network $N(I) = (G(I), d^1, d^2)$.

A *c-coloring* of a graph $G = (V, E)$ can be defined as a mapping

$$C : V \longrightarrow \{0, \dots, c - 1\}$$

such that $C(v) \neq C(v') \quad \forall \{v, v'\} \in E$ (see e.g. [11] for more details). A *binary coloring* of G is a 2-coloring of G .

Given a c -coloring C of $G(I)$ and a color $k \in \{0, \dots, c - 1\}$, each connected component of the subgraph of $G(I)$ generated by the set $\{(i, j) \in P \mid C(i, j) = k\}$ is called a *region*.

Let $\underline{\alpha} = \min \{ \{d_{ij}^1\}_{i \in \{1, \dots, r-1\}, j \in \{1, \dots, s\}}, \{d_{ij}^2\}_{i \in \{1, \dots, r\}, j \in \{1, \dots, s-1\}} \}$ and $\bar{\alpha} = \max \{ \{d_{ij}^1\}_{i \in \{1, \dots, r-1\}, j \in \{1, \dots, s\}}, \{d_{ij}^2\}_{i \in \{1, \dots, r\}, j \in \{1, \dots, s-1\}} \}$. Given a value $\alpha \in [\underline{\alpha}, \bar{\alpha}]$, a way for defining a binary coloring col of $G(I)$ is as follows: Set $col(1, 1) = 0$ and color the rest of the pixels in P from left to right and from up to down, using the following horizontal and vertical rules, respectively:

– Horizontal rule:

$$col(i, j + 1) = \begin{cases} col(i, j) & \text{if } d_{ij}^2 < \alpha \\ 1 - col(i, j) & \text{if } d_{ij}^2 \geq \alpha \end{cases} \quad \forall i \in \{1, \dots, r\}, j \in \{1, \dots, s - 1\}$$

– Vertical rule:

$$col(i + 1, j) = \begin{cases} col(i, j) & \text{if } d_{ij}^1 < \alpha \\ 1 - col(i, j) & \text{if } d_{ij}^1 \geq \alpha \end{cases} \quad \forall i \in \{1, \dots, r - 1\}, j \in \{1, \dots, s\}$$

Obviously, given two colored pixels $(i, j + 1)$ and $(i + 1, j)$ with $i \in \{1, \dots, r - 1\}$ and $j \in \{1, \dots, s - 1\}$, the colors assigned to the pixel $(i + 1, j + 1)$ by means of the horizontal and vertical rules above can be different. When these colors are different, we say that $(i + 1, j + 1)$ is an *inconsistent* pixel.

3 Iterative Binary Coloring Procedure

The relaxed coloring algorithm proposed in [7] for dealing with inconsistent pixels can be outlined as follows:

In the first iteration a value $\alpha_1 \in (\underline{\alpha}, \bar{\alpha})$ is selected, and the binary coloring procedure outlined in Sect. 2 is applied, which results in an assignment of either the color “0” or the color “1” to each pixel. In the second iteration, a value $\alpha_2 \in (\underline{\alpha}, \alpha_1)$ is selected, and the binary coloring procedure is applied, separately, to the subgraph of $G(I)$ generated by those pixels colored as “0” in the first iteration (producing colors “00” and “01”) and to the subgraph of $G(I)$ generated by those pixels colored as “1” in the first iteration (producing colors “10” and “11”). For each successive iteration t , a value $\alpha_t \in (\underline{\alpha}, \alpha_{t-1})$ is selected, and the binary coloring procedure is applied to the subgraphs of $G(I)$ generated by the set of pixels colored with each possible color in the previous iteration. Once all the iterations have been completed, the obtained colors are considered as binary numbers, and the final coloring C is defined by assigning to each pixel the decimal representation of its assigned color.

When an inconsistent pixel is reached, it is arbitrarily colored, but it is isolated so that in the current iteration its inconsistency does not induce inconsistency on its adjacent pixels.

Moreover, a counter for the number of inconsistent pixels is introduced, in such a way that, if its value exceeds a prefixed threshold, the current iteration is aborted.

In order to have a better visualization of the resulting segmentation, all of the pixels of each final region are assigned a new color obtained as the mean of their colors in the original image.

4 Improved Iterative Binary Coloring Procedure

We have introduced some modifications in the procedure described in Sect. 3 that, generally, produce images of similar or better quality and lead to a significant reduction of the number of final regions. The modifications are the following:

- In addition to the movements from left to right and from up to down for the binary coloring definition, we also allow movements from right to left and from down to up, and we redefine the concept of inconsistent pixels according to these new rules.
- Inconsistent pixels are not colored until all the iterations have been completed; then, the final coloring C is defined firstly on the non-inconsistent pixels, and, secondly, on the inconsistent pixels, by considering the Euclidean distances to their adjacent colored pixels.

5 Computational Experience

The procedures described in Sect. 3 and 4 have been applied to an orthoimage of the province of Seville (southern Spain), taken on August 18, 1987, by the LANDSAT 5 Satellite (see Fig. 1). This image was taken with the Thematic Mapper sensor, which has a spatial resolution of 30 m., and it was also considered in 7.



Fig. 1. Original image of Seville



Fig. 2. Segmentation obtained by the original procedure



Fig. 3. Segmentation obtained by the improved procedure

Figures 2 and 3 show, respectively, the segmentations obtained for this image by applying two iterations of the original binary coloring procedure presented in 7 and of our improved procedure. The number of final regions given by the original procedure is 33,639, whereas the improved procedure gives 20,032 regions.

6 Final Comments

The most widely used methods for image segmentation only provide the final picture of the segmentation. On the other hand, a binary coloring procedure for obtaining a hierarchical segmentation of an image was presented in 7, as well as the concept of inconsistencies. The main contribution of this work is to introduce a new way for dealing with inconsistencies which improves previous results providing also hierarchical segmentations of the images.

It is worth noting that, by following the same ideas as in 6,8, the procedure that we have presented could be adapted to obtain hierarchical colorings of fuzzy graphs to be used for image classification, and to obtain a hierarchical clustering technique for networks.

Currently, due to the high computational effort required, there are not many algorithms that provide a hierarchical clustering of a set of pixels. Our hierarchical segmentation algorithm shows the evolution of the segmentation process iteratively. Therefore, in addition to the hierarchical clustering of an image, this algorithm provides an easy way to build the *Border* and *Non-Border* fuzzy classes of an image.

Some authors stress on the absolute need of new methodologies for assessing the performance of clustering and classification algorithms when the data show a certain structure and the output is not given in the classical way (see e.g. 5). Several problems are still open within this framework, as, for example, the way in which two dendrograms are compared.

Acknowledgments. This research has been partially supported by the Government of Spain, Grant TIN2009-07901.

References

1. Boskovitz, V., Guterman, H.: An Adaptive Neuro-Fuzzy System for Automatic Image Segmentation and Edge Detection. *IEEE Trans. Fuzzy Syst.* 10, 247–262 (2002)
2. Carson, C., Belongie, S., Greenspan, H., Malik, J.: Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 1026–1038 (2002)
3. Cheng, H.D., Li, J.: Fuzzy Homogeneity and Scale-Space Approach to Color Image Segmentation. *Pattern Recognit.* 36, 1545–1562 (2003)
4. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient Graph-Based Image Segmentation. *Int. J. Comp. Vis.* 59, 167–181 (2004)
5. Gómez, D., Montero, J., Biging, G.: Accuracy Statistics for Judging Soft Classification. *Int. J. Remote Sens.* 29, 693–709 (2008)
6. Gómez, D., Montero, J., Yáñez, J.: A Coloring Fuzzy Graph Approach for Image Segmentation. *Inf. Sci.* 176, 3645–3657 (2006)
7. Gómez, D., Montero, J., Yáñez, J., Poidomani, C.: A Graph Coloring Approach for Image Segmentation. *Omega* 35, 173–183 (2007)
8. Gómez, D., Yáñez, J., Montero, J.: Bi-Criteria Clustering in Networks (submitted, 2011)
9. Grady, L.: Random Walks for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 1768–1783 (2006)
10. Grady, L., Schwartz, E.L.: Isoperimetric Graph Partitioning for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 469–475 (2006)
11. Gross, J., Yellen, J.: *Graph Theory and its Applications*. CRC Press, Boca Raton (1999)
12. Hung, W.L., Chang, Y.C., Lee, E.S.: Weight Selection in W-K-Means Algorithm with an Application in Color Image Segmentation. *Comput. Math. Appl.* 62, 668–676 (2011)
13. Lermé, N., Létocart, L., Malgouyres, F.: Reduced Graphs for Min-Cut/Max-Flow Approaches in Image Segmentation. *Electron. Notes Discret. Math.* 37, 63–68 (2011)
14. Liu, J., Yang, Y.H.: Multiresolution Color Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 16, 689–700 (1994)
15. Malik, J., Belongie, S., Leung, T., Shi, J.: Contour and Texture Analysis for Image Segmentation. *Int. J. Comp. Vis.* 43, 7–27 (2001)
16. Martín, J.A., Montero, J., Yáñez, J., Gómez, D.: A Divisive Hierarchical k-Means Based Algorithm for Image Segmentation. In: *Proc. ISKE 2010*, pp. 300–304 (2010)
17. Shi, J., Malik, J.: Normalized Cuts and Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 22, 888–905 (2000)
18. Tobias, O.J., Seara, R.: Image Segmentation by Histogram Thresholding Using Fuzzy Sets. *IEEE Trans. Image Process.* 11, 1457–1465 (2002)
19. Wu, Z., Leahy, R.: An Optimal Graph Theoretic Approach to Data Clustering: Theory and its Application to Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 15, 1101–1113 (1993)
20. Yáñez, J., Muñoz, S., Montero, J.: Graph Coloring Inconsistencies in Image Segmentation. In: *Proc. FLINS 2008*, pp. 435–440 (2008)

Moving Objects Detection Using Adaptive Region-Based Background Model in Dynamic Scenes

Lin Gao, Yong Fan, Niannian Chen, Yufeng Li, and Xiaorong Li

Department of Computing Science and Technology, Southwest University of Science and Technology, 621010 Mian Yang, China

Abstract. Moving object detection plays an important role in video surveillance, yet in dynamic scenes it is still a challenging problem. In this work, we develop an efficient algorithm to handling complex dynamic backgrounds by using an adaptive region-based background model. Firstly, the initial background image is partitioned using image over-segmentation methods. Then the input frame is partitioned to image regions according to the obtained partition manner. Features of the image regions are used to construct adaptive mixture Gaussian models. When the background model is updating, the number of component of the mixture Gaussian models is selected adaptively based on the activity level of features. A coarse-to-fine strategy is designed to detect the moving object. The foreground and background are distinguished gradually in region-level and pixel-level through the built background model. Experimental results show that the algorithm proposed in this paper can detect moving objects quickly and effectively.

Keywords: moving object detection, background modeling, mixture of Gaussians, image over-segmentation.

1 Introduction

Moving objects detection in video sequence is to separate the regions corresponding to the moving objects from the background in the video sequence. The detection results are usually used as interested region for high-level visual processing, such as object tracking, object recognition and motion analysis etc. Under the condition of stationary cameras, a common approach is to perform background subtraction [1-7], which identifies moving objects from the video frame that differs significantly from a background model, thus the effectiveness of the detection methods relies on the background model. However, many traditional background models typically fail in dealing with the dynamic background changes, such as lighting changes, moving elements of the scene, etc. In order to solve these problems, models has become more complicated, which brings about high computation complexity, resulting in contradiction with the real-time demand of the system. Currently, most of the models treat pixel as basic processing unit, and the most typical methods are Gaussian mixture background model proposed by Stauffer[1] and its improved algorithms[8-11]. Because of the gigantic amount of pixels, the amount of computation is too big to satisfy the real-time demand, besides, since there are certain spatial correlations between adjacent pixels,

redundant information contained in these correlations is ignored, which results in waste of huge computer resource. Considering this problem, Sheikh[12] described neighborhood relationship between pixels using Markov Random Field, but it was difficult to solve MRF as well as big calculations.

To deal with the shortage of pixel-based background model, region-based algorithms were proposed, which usually divide an image into blocks and calculate region-specific features. These methods enable subtraction of the amount of processing units, thus reducing the complexity of the algorithms, but accuracy goes down as well. In addition, the way of how to divide into regions is also a key factor to the performance of the model. Many approaches partition the image region into rectangular blocks because of its convenience [13-15]. However, the blocks divided in this way lack linguistic correlations with the objects in the scene to some extent, disabling accurate description of the background structure and the spatial correlations between pixels, which lowers the precision of the background model. Apparently, if the spatial correlations between adjacent pixels are made full use of, and pixel-level processing accuracy can be maintained, highly effective background model would be obtained.

Based on Gaussian mixture background model, we proposed moving objects detection algorithm based on an adaptive region-based background model. Firstly, the initial background image is partitioned using image over-segmentation methods. Then the input frame is partitioned to image regions according to the obtained partition manner. Features of the regions are extracted to construct adaptive mixture Gaussian model. When the background model is updating, the number of component of the mixture Gaussian is selected adaptively based on the activity level of features. A coarse-to-fine strategy is designed to detect the moving object. The foreground and background are distinguished gradually in region-level and pixel-level through the built background model.

2 Image Region Segmentation Based on Image Over-Segmentation

Image over-segmentation is an image preprocessing method which has been developed recently [16-18]. It divides images into regions that are uniform and adjacent to each other (which are also called “superpixels”). These regions are able to maintain structural information in interested scale, as shown in Figure 1. The reasons why image over-segmentation is treated as image preprocessing are based on the following two considerations: First, pixel itself is not a natural entity, but a product of discretization of images, it therefore has no instructive meaning in image analysis. Second, the amount of pixels is usually huge, so pixel-level optimization algorithms are difficult to put into practice because of big calculations, while the complexity of computation is largely reduced when using regions as basic processing units after image over-segmentation.

Usual over-segmentation methods includes Normalized cuts algorithm[19], feature space analysis based on Mean Shift[20]. In this paper, we use Superpixels Lattices method[18]. Comparing to other algorithms, this algorithm has the following feature: the blocks after segmentation maintain the same grid topology structure shape as what it has when the image is expressed by the original pixels, as is shown in the third column of figure 1. This topology type offers convenience for processing afterwards.

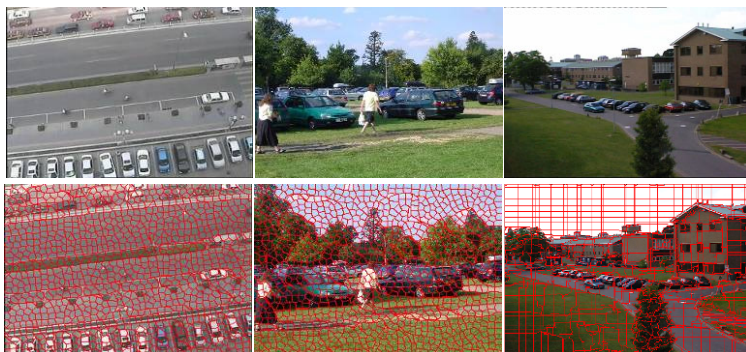


Fig. 1. Image segmentation processing (original images in the first row, and corresponding results of over-segmentation in the second row)

3 Gaussian Mixture Background Model Based on Region Features

By means of image over-segmentation, the background of the scene is divided into many regions (ie. superpixels), each of which has certain semantics and has comparatively stable structures. For example, pixels in the same region have close values, in addition, when lighting condition changes, the changes of the color value of the pixels are consistent. The benefits of background modeling based on superpixel are: it can overcome noise effect, adapt background disturbance; it largely reduce the redundancy computing and improve the computation speed.

3.1 Extraction of Region Features

Because superpixels are image regions obtained by image over-segmentation, the color value of every pixel converges inside the superpixel which represents the background. In other words, the color distribution of background regions highly centralizes in a small section of the histogram. Comparatively, variety of foreground objects, whose color and texture feature widely distribute in the feature space, has small possibility to have the consistent character distribution with that of background in superpixel region. Therefore, a small amount of feature vectors to describe the background is enough to distinguish foreground with background.

We use moment feature to describe the color distribution in each region. Since color distribution information mainly concentrate in low order moment, hence we use Mean m_1 , Variance m_2 , and Skewness m_3 to describe region features. Assumption that we have obtained an $M \times N$ collection of superpixels S , S_i stands for the i th superpixel, \mathbf{v}_i stands for the feature vector extracted from S_i . Then $\mathbf{v}_i = (S_i(m_1), S_i(m_2), S_i(m_3))$.

3.2 Background Modeling and Update

The change of scene will finally be reflected by the change of superpixel feature vector \mathbf{v} , thus we can utilize feature vector for background modeling, by which we can tell that the superpixel belongs to background or foreground. The same as the background pixel color distribution, the distribution of vector \mathbf{v} may be in multimode as well. Marked in Figure 2(a) is a superpixel S_a of a shaking tree branch. Figure 2(b) shows the S_a 's feature values in consecutive 500 frames, where we can see the feature value is in multimodal distribution. The two ellipsoids in the figure indicate that the distribution can be fitted by two Gaussian mixtures.

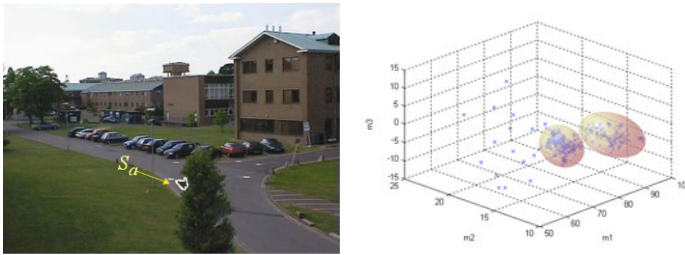


Fig. 2. The distribution of Superpixel's feature values in time sequence

Given a time period, the i th superpixel's feature vectors constitute a collection $\{\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,t}\}$. According to historical value that the feature vector has ever been, we use K Gaussian mixture models to stand for the probability distribution of feature, namely the probability of superpixel i 's $\mathbf{v}_{i,t}$ at time t is

$$p(\mathbf{v}_{i,t}) = \sum_{j=1}^K w_{i,t}^j \eta(\mathbf{v}_{i,t}; \theta_{i,t}^j) \tag{1}$$

Where $w_{i,t}^j$ is the weight of the j th Gaussian component in the model, $\eta(\mathbf{v}_{i,t}; \theta_{j,t}^j)$ is the probability density function of the j th Gaussian component.

$$\begin{aligned} \eta(\mathbf{v}_{i,t}; \theta_{i,t}^j) &= \eta(\mathbf{v}_{i,t}; \mu_{i,t}^j, \Sigma_{i,t}^j) \\ &= \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_{i,t}^j|^{\frac{1}{2}}} \exp\left[-\frac{1}{2} (\mathbf{v}_{i,t} - \mu_{i,t}^j)^T (\Sigma_{i,t}^j)^{-1} (\mathbf{v}_{i,t} - \mu_{i,t}^j)\right] \end{aligned} \tag{2}$$

Where $\mu_{i,t}^j$ is the mean value, $\Sigma_{i,t}^j$ is the covariance matrix, and D is the dimension of $\mathbf{v}_{i,t}$.

Since the scene is usually dynamic, such as the change of lighting condition, the movement of background objects, etc, we need to update the model accordingly. The update of Gaussian mixture background model refers to the update of the weight and other parameters of each Gaussian component. And the specific procedures are as follows: First, sort the K Gaussian components by $w_{i,t}^j / \sigma_{i,t}^j$ in descending order. Then use the superpixel's current feature $\mathbf{v}_{i,t}$ to compare each of the Gaussian component one by one, if the distance between $\mathbf{v}_{i,t}$ and the j th component's mean value $\mu_{i,t}^j$ is smaller than τ times standard deviation $\sigma_{i,t}^j$, ie. $\|\mathbf{v}_{i,t} - \mu_{i,t}^j\| \leq \tau \sigma_{i,t}^j$, then consider they match (usually τ ranges from 2.5~3.5). Gaussian component that matches will be updated. We use the equation proposed by Stauffer and Grimson [1]:

$$w_{i,t}^j = (1 - \alpha)w_{i,t-1}^j + \alpha \cdot M_{i,t-1}^j \tag{3}$$

$$\mu_{i,t}^j = (1 - \rho_{i,t}^j)\mu_{i,t-1}^j + \rho_{i,t}^j \cdot \mathbf{v}_{i,t} \tag{4}$$

$$\sigma_{i,t}^j = (1 - \rho_{i,t}^j)(\sigma_{i,t-1}^j)^2 + \rho_{i,t}^j(\mathbf{v}_{i,t} - \mu_{i,t-1}^j)^T(\mathbf{v}_{i,t} - \mu_{i,t-1}^j) \tag{5}$$

where

$$\rho_{i,t}^j = \eta(\mathbf{v}_{i,t} \mid \mu_{i,t}^j, \sigma_{i,t}^j) \tag{6}$$

where α is the learning rate of the model, and $0 < \alpha < 1$, which reflects the speed of background updating. When α is larger, the change of the background will be rapidly reflected in the model, on the contrary, when the α is small, parameters of the model will be updated slowly as background changes. $M_{i,t-1}^j$ is the matching operator, $M_{i,t-1}^j = 1$ when component matches $\mathbf{v}_{i,t}$, $M_{i,t-1}^j = 0$ otherwise.

In Gaussian mixture model, we deem that Gaussian components with larger weight and smaller variance are generated from background, namely, in the order by $w_{i,t}^j / \sigma_{i,t}^j$, we take the first B_i Gaussian components as background distributions. B_i 's value is:

$$B_i = \arg \min_b \left(\sum_{j=1}^b w_{i,t}^j > T \right) \tag{7}$$

where T is the threshold for the minimum proportion of the date that should be accounted for by the background.

3.3 The Adaptive Amount of Gaussian Components

Usual Gaussian mixture background models often use fixed number of components. Dynamic regions need more components to describe its multimodal distribution,

while regarding to static regions, one Gaussian is enough. According to statistics, most background regions in a video sequence are in a static state. Thus, in these regions, if still using more Gaussian components, it will cause waste of computing resource. Therefore, we can use different amount of Gaussian components in different time periods or different regions, hence improving the efficiency of the background model.

We can conclude, according to the updating rules of Gaussian mixture model, in dynamic background regions, Gaussian component that matches the background occurs alternately, while in static background regions, one can keep in matching state for a long time, other components idle. Therefore, we adjust each superpixel's Gaussian components according to the following strategy.

Although the weight of Gaussian component can more or less reflects its idle extent, where smaller weight corresponds to higher idle extent, however, with regard to background regions that in multimode, some Gaussian components belonging to backgrounds also have smaller weight. So it's not reliable to tell whether it's in long idle state by merely judging its weight. For solving this problem, we add an attribute a_j to each component, measure how long it lasts for staying matched with the scene.

When it lasts longer, the value of a_j is larger, or else smaller. When judging, one should consider both the last degree a_j and $w_{i,t}^j / \sigma_{i,t}^j$. For a certain superpixel S_i , the adjustment procedures of the amount of Gaussian components in the model are as follows:

- 1) When initializing background model, set only one Gaussian component in Gaussian mixture model, and initialize $a_j = a_{init}$. Gradually increase the amount of Gaussian components as model updates.
- 2) When the j th Gaussian component of superpixel i matches the scene, add a_j by 1 (if a_j becomes larger than maximum Ta_{max} after addition, then $a_j = Ta_{max}$), and subtract a_j by 1 of other components (if a_j is smaller than 0 after subtraction, then $a_j = 0$). if the Gaussian mixture model of the superpixel cannot match the scene, decrease all Gaussian component's last degree, then create a new component using the superpixel's current value as mean value. If the amount of Gaussian components reaches the maximum value K_{max} , then the new created Gaussian component replaces the one in the bottom of the list, else, add this new Gaussian component to the Gaussian mixture model.
- 3) After the background model learning the current scene, check every Gaussian component of superpixel i , if there are more than 1 components, and the last degree a_j of the Gaussian component which is the last one in the order equals 0, then delete it.

4 Moving Object Detection

4.1 Region Level Detection

The usual method to distinguish between foreground and background is to compare the current frame with the background, if the difference of a certain pixel is larger than the threshold value, then it's supposed to be foreground pixel, otherwise background pixel. Different from this method, in method based on Gaussian mixture background model, when the model finishes its updating, the distinction between background and foreground is completed too. As for the superpixel in this paper, when the correspondent Gaussian model finishes its updating, according to formula (7), the first B components are used to describe background, and the others are foreground distribution. If the feature of superpixel matches the first B component, judge it background, otherwise foreground.

4.2 Pixel Level Detection

In this work, we use two pixel features for classification: color and edge feature. First take into account the color feature. If the granularity of the superpixel is small enough, the pixel color of the background region will be concentrated. Compare it with the mean color value of the background region, if the difference is big enough, consider it foreground pixel. The color difference is defined as follows:

$$d_{c,t}(x, y) = \sum_{j=1}^B |\mu_{i,t}^j(\mathbf{1}) - C_t(x, y)| w_{i,t}^j \quad (8)$$

where $C_t(x, y)$ is the color of pixel $I_t(x, y)$, the corresponding superpixel is $S_{i,t}$, $\eta(\mathbf{v}_{i,t}; \mu_{i,t}^j, \Sigma_{i,t}^j)$ is the j th Gaussian component of $S_{i,t}$, $\mu_{i,t}^j(\mathbf{1})$ is the first order moment of the region color, $w_{i,t}^j$ is the weight of Gaussian component. If $d_{c,t}(x, y)$ is larger than the threshold τ_c , then consider pixel $I_t(x, y)$ is foreground pixel, otherwise background pixel.

We can tell whether the edge belongs to the background or not by judging from its intensity. Because the shape of the superpixel is obtained by division according to the edge of background, inside the region is usually smooth, thus only edges with weaker intensity appears in these regions. By contrast, foreground objects may have strong-intensity edges. Thus, with regard to pixels in a superpixel, the stronger intensity the edges have, the greater probability it belongs to foreground. The intensity of the pixel's edge can be measured by its gradient $\|\text{grad}(x, y)\|$, namely the larger the gradient is, the more possible the pixel belongs to foreground. Firstly, use Sobel operator to approximate the first-order derivative in the horizontal direction and the vertical direction, and then calculate the gradient $\|\text{grad}(x, y)\|$.

Based on the analysis above, considering color and edge feature, a pixel is discriminated by:

$$I_f(x, y) \in \begin{cases} \text{Foreground} & \text{if } (d_{c,t}(x, y) > \tau_c) \vee (\|\text{grad}(x, y)\| > \tau_e) \\ \text{Background} & \text{otherwise} \end{cases} \quad (9)$$

where τ_e is threshold value of edge intensity.

Figure 3 demonstrates an example using the classification strategy mentioned in this section, (a) is the original image, (b) is the detection result obtained by manual segmentation, (c) is region level detection result, (d) is the pixel level detection result. From Figure 3, we can see that detection precision can be improved through the coarse-to-fine process.

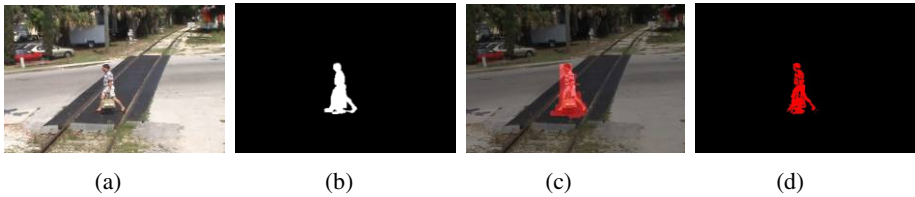


Fig. 3. Using coarse-to-fine strategy for moving object detection

5 Experimental Results

We evaluated the performance of the proposed algorithm on several representative sequences, and compared our algorithm with traditional Gaussian mixture background algorithm[1] (MoG) and background modeling method proposed by Sheikh in [12]. All the algorithms are programmed with VC6.0. The hardware platform : 3.0GHz Pentium 4, 1GB RAM. In these experiments, the parameters are pre-set as follow: image over-segmentation parameter $M = 45$, $N = 50$, background modeling parameter $K_{\max} = 5$, $T = 0.6$, $\alpha = 0.002$, $\tau = 2.5$, $a_{\text{init}} = 30$, $Ta_{\max} = 125$, foreground object detection parameter $\tau_c = 60$, $\tau_e = 0.4$. Parameters of MoG are the same with that of ours, and parameters of Sheikh algorithm refers to [12].

Figure 4 shows the detection results of the contrast experiments, in which the test sequences are all public data sets for objects detection. The first row shows the original testing frames; the second row shows the ground truth image obtained by manual segmentation, used for quantitative comparison and analysis of detection results; the third row demonstrates the detection result using MoG algorithm; the fourth row shows the result obtained by Sheikh algorithm; and the fifth row is the detection result using our method.

In WavingTress sequence, because of the dynamic changes of waving trees, MoG algorithm cannot adapt to the dynamic background, resulting in comparatively more false positives. Sheikh algorithm and our algorithm, since both of them has taken the neighborhood relationships of pixels into account, can both eliminate the effect of irregular changes, so that the false positives are far less than that of MoG's. The difficulty in TimeOfDay is the apparent change of the lighting condition, and the

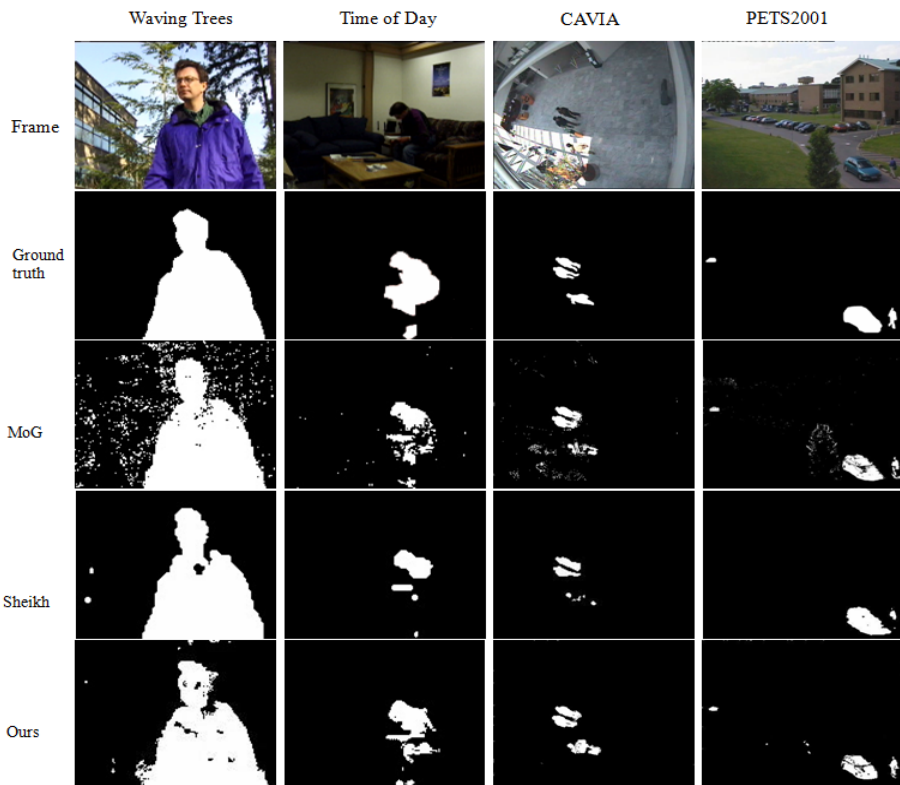


Fig. 4. The detection results on some public datasets

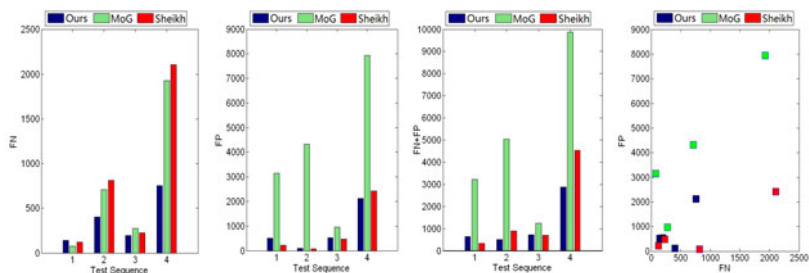


Fig. 5. The quantitative comparison of the detection results

Table 1. Processing speed of three algorithms (fps)

Sequence	MoG algorithm	Sheikh algorithm	Our algorithm
Waving Trees (120×160)	52.5	49.2	91.3
Time of Day (120×160)	51.1	47.6	89.7
CAVIAR (384×288)	13.8	12.9	49.7
PETS2001 (768×576)	7.8	6.4	21.3

foreground objects are hard to distinguish from background in insufficient light. Seen from the detection results, MoG still has many false positives and Sheikh algorithm has serious problem in false negatives. Our algorithm has lower accuracy than Sheikh, but the false positives are far more less. As for CAVIAR, the disturbance mainly comes from the disordered background objects, and the noise caused by the disturbance will affect the accuracy of segmentation. The results show that our algorithm is better than the other two. In PETS2001, the noise of the camera sensor and the waving tree branches all cause disturbance to the establishment and the update of the background model. MoG has serious false positives problem, which mainly focus on the edge of the tree branches. Sheikh algorithm occurs false negatives, and our algorithm can basically overcome the disturbance. As for the detection precision, it has the similar results with the other two algorithms.

Figure 5 demonstrates the quantitative contrast results. We use the method proposed in [21] to assess the effectiveness of the methods, and the evaluation indexes are include False Positives(FP) and False negatives(FN). FP and FN are all obtained by contrasting with the ground truth image. As can be seen, in all the sequences, the FP and FN of our algorithm keep in a lower level, while Sheikh algorithm has comparatively higher FN. As for overall indicator, our algorithm is better than MoG, and is similar to Sheikh algorithm.

Table 1 shows the time spent by the three algorithms on processing the four sequences. MoG algorithm models every individual pixel so that its speed is comparatively slower. In Sheikh algorithm, the most likely configuration of MRF is obtained using the maximum a posteriori probability estimate, which contains huge computation, costing more time. Our algorithm uses region level background model, which reduces the number of processing units to a large extent, in spite of going through a two-level processing from region-level to pixel-level, it still has apparent improvement of computing speed than other methods.

6 Conclusion

In this paper, we propose a moving objects detection algorithm based on an adaptive background model. In order to improve the modeling effectiveness of the background model, we use image over-segmentation technique to divide background into regions, and then use Gaussian mixture model for each individual region. The proposed background model is able to adjust the amount of Gaussian component according to the activity degree of the scene change, thus achieving further improvement of the effectiveness of the model. When processing the moving objects detection, we use a coarse-to-fine detection strategy that is from region-level to pixel-level, thus ensuring the robustness as well as improving the precision of target detection. The experimental results demonstrated that the proposed algorithm is fast and reliable and meets the requirement of real-time system.

References

1. Stauffer, C., Grimson, W.E.L.: Learning patterns of activity using real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(8), 747–757 (2000)

2. Elgammal, A.M., Duraiswami, R., Harwood, D., Davis, L.S.: Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proc. IEEE* 90, 1151–1163 (2002)
3. Hu, W., Gong, H., Zhu, S.-C., Wang, Y.: An integrated background model for video surveillance based on primal sketch and 3D scene geometry. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2008)
4. Kim, K., Chalidabhongse, T.H., Harwood, D., Davis, L.S.: Background modeling and subtraction by codebook construction. In: *Proc. IEEE International Conference on Image Processing*, pp. 3061–3064 (2004)
5. Liao, S., Zhao, G., Kellokumpu, V., Pietikainen, M., Li, S.Z.: Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, p. 8 (2010)
6. Toyama, K., Krumm, J., Brumitt, B., Meyers, B.: Wallflower: Principles and practice of background maintenance. In: *Proc. International Conference on Computer Vision*, vol. 1, pp. 255–261 (1999)
7. Yao, J., Odobez, J.-M.: Multi-layer background subtraction based on color and texture. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8 (2008)
8. Zivkovic, Z., van der Heijden, F.: Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters* 27(7), 773–780 (2006)
9. Lee, D.S.: Effective gaussian mixture learning for video background subtraction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 827–832 (2005)
10. KaewTraKulPong, P., Bowden, R.: An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection. In: *The 2nd European Workshop on Advanced Video-based Surveillance Systems* (2001)
11. Power, P.W., Schoonees, J.A.: Understanding Background Mixture Models for Foreground Segmentation. In: *Proceedings Image and Vision Computing, New Zealand* (2002)
12. Sheikh, Y., Shah, M.: Bayesian Modeling of Dynamic Scenes for Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(11) (2005)
13. Chen, Y.-T., Chen, C.-S., Huang, C.-R., Hung, Y.-P.: Efficient hierarchical method for background subtraction. *Pattern Recognition* 40(10), 2706–2715 (2006)
14. Alexandropoulos, T., Loumos, V., Kayafas, E.: A block-based clustering technique for real time object detection on a static background. In: *2nd International IEEE Conference on Intelligent Systems*, pp. 169–173 (2004)
15. Russell, D., Gong, S.: A highly efficient block-based dynamic background model. In: *Advanced Video and Signal Based Surveillance*, pp. 417–422 (2005)
16. Ren, X., Malik, J.: Learning a Classification Model for Segmentation. In: *Proceedings of the International Conference of Computer Vision* (2003)
17. Li, Y., Sun, J., Tang, C.K., Shum, H.Y.: Lazy snapping. *ACM Transactions on Graphics (TOG)* 23(3), 303–308 (2004)
18. Moore, A.P., Prince, S.J.D., Warrell, J., Mohammed, U., Jones, G.: Superpixel lattices (2008)
19. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence* 22(8), 888–905 (2000)
20. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 603–619 (2002)
21. Heikkila, M., Pietikainen, M.: A texture-based method for modeling the background and detecting moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(4), 657–662 (2006)

New Shape-from-Shading Method with Near-Scene Point Lighting Source Condition

Lei Yang¹, Shiwei Ma¹, and Bo Tian²

¹ Shanghai Key Laboratory of Power Station Automation Technology,
School of Mechatronics Engineering and Automation, Shanghai University,
Yan-chang Road 149, Shanghai, 200072, China

² School of Information Management & Engineering,
Shanghai University of Finance and Economic, Guo-ding Road 777, Shanghai, 200433, China
yangyoungya@gmail.com

Abstract. Imaging model is one of key factors of shape-from-shading (SFS) which reconstruct 3-D shape of objectives from only one image. Most existing SFS methods are based on orthogonal projection. But perspective projection is more accurate than orthogonal projection to simulate the imaging processes of cameras. In this paper, new SFS method under perspective projection with near-scene point lighting source condition is proposed. Imaging model under perspective projection with near-scene lighting source condition is formulated firstly. And the influence of the distance from lighting source to the surface is also considered in the reflectance map equation. Then a set of static Hamilton-Jacobi (H-J) equation is given by the reflectance map equation. The SFS problem is further formulated as viscosity solution of the H-J equation. And the Lax-Friedrichs fast sweeping numerical scheme based on the vanishing viscosity approximation theory is used to solve the H-J equation. At last, experiments on synthetic and real images are performed. Experimental results illustrated the efficiency of the proposed method.

Keywords: Shape-from-shading, perspective projection, point lighting source, static Hamilton-Jacobi equation, viscosity solution.

1 Introduction

Shape-from-shading (SFS) is an approach of 3-D shape reconstruction from only single image, which is a one of the fundamental problems in computer vision [1-3]. The brightness of image mainly depends on four factors: the orientation of light source, the location of camera, the orientation of object and the reflectance properties of the surface [4]. SFS methods reconstruct the 3-D shape of objects from the shading variation in 2-D image, which is the process that solves the reflectance map equation at each imaged point. The variation of brightness is used to estimate the orientation of surface and then calculate the height of objects. The mathematic formulate of SFS and first SFS algorithm is proposed by Horn [1]. And SFS is widely applied in industry automatic inspection, face recognition, terrain analysis, and so on [5-8].

The original SFS methods proposed by Horn minimize the total error function consisting of one or several terms of the brightness constraint, the smoothness constraint by using the principle of variations [1]. After SFS problem is formulated, different computational techniques have been introduced into SFS field. Horn's SFS algorithms belong to minimization methods [2-3]. Linear reflectance map method is proposed in 1990s [2]. At that time, heuristic approach and other computing techniques are researched [9-13]. Zhang categorized the traditional SFS approach into four categories: minimization, propagation, local, and linear approaches [2]. After that, neural networks methods have also been employed in SFS [13]. Recently, statistical SFS approaches have been proposed [12]. Other SFS reconstruction method such as varying reflectance model and Color SFS was also brought forward [4]. Durou surveyed numerical SFS methods [3]. Perspective SFS methods have been paid more attention to in the last ten years [14-16]. Numerical algorithms for perspective SFS were surveyed [17]. SFS can also be formulated as optimal control problems [18-19]. And 3-D reconstruction techniques using fusing strategies which compose SFS method and other approaches as Photometric Stereo, series images or others are investigated [20-23]. As for the applications of SFS, Horn discussed how to reconstruct 3-D surface of the Moon [1]. Song proposed new defecting detection method using SFS [6]. Bors analyzed radar SFS problem [8]. Human face recognition based on SFS was also proposed [24]. Several researchers reconstructed 3-D shapes of medical images using SFS method [15-16].

Among these, partial-differential-equations (PDEs) based SFS algorithm is an important class of reconstruction method. Characteristics-lines method is the earliest one [2]. Using programming theory, SFS is solved as viscosity solutions of static Hamilton-Jacobi (H-J) equations [16, 22]. There are mainly two classes of numerical methods for solving static H-J equations. The first class of numerical methods is based on reformulating the equations into suitable time-dependent problems [16, 22-23]. Heat equation based SFS [23] proposed recently belongs to this class. The other class of numerical methods is to formulate the problem as a stationary boundary value problem and discretize the problem into a set of nonlinear algebra numerical equations [16, 24]. Such methods are the fast sweeping method [16], and Level set propagation [25]. As we know, SFS methods essentially are solving some PDEs derived from reflectance map equations which are determined by the reflectance model and certain constraints conditions. So suitable imaging models and effective reconstruction algorithms are two main directions researched in SFS field. Most existing SFS methods are based on orthogonal projection in which the light source was assumed to be located at infinite distance [1-4]. But as we know, perspective projection is more accurate than orthogonal projection to simulate the imaging processes. In the last ten years, perspective projection was gradually adopted in SFS [15-17]. In this paper, a new SFS method under perspective projection with near-scene point lighting source condition is proposed. Remainders of the paper are organized as follows. Imaging model is formulated in section 2. And a set of static Hamilton-Jacobi (H-J) equation is given by the reflectance map equation. And the SFS problem is formulated as viscosity solution of the H-J equation. In section 3, the Lax-Friedrichs fast sweeping numerical scheme based on the vanishing viscosity approximation theory is adopted to solve the SFS H-J equation. Experimental 3-D reconstruction results of synthetic and real images are reported in section 4. At last, conclusions are drawn in the fifth section.

2 The Proposed Perspective Reflectance Model

It is almost impossible to establish unified imaging model in SFS methods because of different imaging conditions in practice. Horn used the following Lambertian reflectance equation to solve the 3-D shape of the Moon surface.

$$E = \rho \cos \theta \quad (1)$$

where E is the captured image, θ is angle of surface normal \vec{n} and lighting source vector \vec{s} , and ρ is constant albedo of Lambertian reflectance surface. If following assumptions are made: the lighting source is located at infinite distance with a known direction $\vec{s} = (p_0, q_0, -1)$, surface normal \vec{n} is denoted by $(p, q, -1)$, and orthogonal projection is used to simulate the imaging process, following reflectance map equation is established.

$$I = \frac{E - E_{\min}}{E_{\max} - E_{\min}} = \frac{\vec{n} \cdot \vec{s}}{|\vec{n}| |\vec{s}|} = \frac{pp_0 + qq_0 + 1}{\sqrt{p^2 + q^2 + 1} \sqrt{p_0^2 + q_0^2 + 1}} \quad (2)$$

where I is the normalized image, E_{\max} and E_{\min} denote the maximum and minimum grey values of the captured image E , and $(p, q) = (\partial z / \partial x, \partial z / \partial y)$, $z = z(x, y)$ is the surface in scene. After Horn proposed the reflectance map equation, different mathematic techniques were adopted to solve the equation [2].

As for near-scene imaging situation such as medical endoscope image, following assumptions are reasonable: (H1) Perspective projection is used (because image is formulated in near-scene), (H2) Point light source is located at the optical center (because light source is near the camera lens as for endoscope imaging process), (H3) And the surface is Lambertian with constant albedo. Then the reflectance map equation of Lambertian reflectance surface under irradiation of near scene point lighting source is as [16]

$$I = \rho \frac{\vec{n} \cdot \vec{s}}{r^2 * |\vec{n}| |\vec{s}|} \quad (3)$$

where r is the distance from object surface to lighting source. We will formulate the reflectance map equation Eq.(3) under perspective projection. The surface S in scene is defined as $z = z(x, y)$, and pixel is denoted as (u, v) at imaging plane $z = f$, where f is focal length. The perspective projection equations is

$$\frac{x}{u} = \frac{y}{v} = \frac{z}{f} \quad (4)$$

Form Eq.(4), we have

$$\begin{cases} x = u \frac{z}{f} \\ y = v \frac{z}{f} \end{cases} \tag{5}$$

The surface S is parameterized in imaging plane such as

$$(x, y, z) = z(u, v) * (\frac{u}{f}, \frac{v}{f}, 1) \tag{6}$$

By using the difference geometry theory, the normal of the surface is given by

$$\begin{aligned} \vec{n}(u, v) = \vec{S}_u \times \vec{S}_v &= \begin{vmatrix} i & j & k \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \end{vmatrix} = \begin{vmatrix} i & j & k \\ \frac{z_u u + z}{f} & \frac{z_u v}{f} & z_u \\ \frac{uz_v}{f} & \frac{z_v v + z}{f} & z_v \end{vmatrix} \\ &= (-f \frac{\partial z}{\partial u}, -f \frac{\partial z}{\partial v}, u \frac{\partial z}{\partial u} + v \frac{\partial z}{\partial v} + z(u, v)) * \frac{z(u, v)}{f^2} \end{aligned} \tag{7}$$

And the unit normal of the surface is

$$\vec{n} = \frac{(-fz_u, -fz_v, uz_u + vz_v + z)}{\sqrt{(uz_u + vz_v + z)^2 + f^2(z_u^2 + z_v^2)}} \tag{8}$$

where $(z_u, z_v) = (\partial z / \partial u, \partial z / \partial v)$. Because the lighting source is located approximately at optical center $(0, 0, 0)$, the direction vector of lighting source at surface (x, y, z) is $\vec{s} = (u, v, f)$, and the unit normal of lighting source is

$$\vec{s} = \frac{(u, v, f)}{\sqrt{u^2 + v^2 + f^2}} \tag{9}$$

Because the point (x, y, z) of surface S is parameterized as Eq.(6), the distance from lighting source to the point (x, y, z) of surface S is

$$r = \frac{z}{f} \sqrt{u^2 + v^2 + f^2} \tag{10}$$

So the reflectance map equation from Eq.(3) is as

$$I(u, v) = \sigma \frac{\vec{n} \cdot \vec{s}}{r^2} = \frac{\sigma f^3}{z(\sqrt{u^2 + v^2 + f^2})^3 \sqrt{(uz_u + vz_v + z)^2 + f^2(z_u^2 + z_v^2)}} \quad (11)$$

When the reconstructed surface is in the imaging plane, normalized grey values of at (0, 0) of the image plane is maximum. From Eq.(11), $(u, v) = (0, 0)$, $(z_u, z_v) = (0, 0)$, $I(0, 0) = 1$ and $z = f$, we have $\sigma = f^2$. And the normalized reflectance map equation from Eq.(11) is

$$I(u, v) = \frac{f^5}{z(\sqrt{u^2 + v^2 + f^2})^3 \sqrt{(uz_u + vz_v + z)^2 + f^2(z_u^2 + z_v^2)}} \quad (12)$$

Equation (13) can be further written as

$$I(u, v) = \frac{f^5}{z^2(\sqrt{u^2 + v^2 + f^2})^3 \sqrt{\left(u \frac{\partial \ln z}{\partial u} + v \frac{\partial \ln z}{\partial v} + 1\right)^2 + \left(f \frac{\partial \ln z}{\partial u}\right)^2 + \left(f \frac{\partial \ln z}{\partial v}\right)^2}} \quad (13)$$

Let $Z(u, v) = \ln z(u, v)$, and we denote $z(u, v) > 0$, then we have

$$H(u, v, \frac{\partial Z}{\partial u}, \frac{\partial Z}{\partial v}) = I(u, v) \sqrt{\left(f \frac{\partial Z}{\partial u}\right)^2 + \left(f \frac{\partial Z}{\partial v}\right)^2 + \left(u \frac{\partial Z}{\partial u} + v \frac{\partial Z}{\partial v} + 1\right)^2} - \frac{f^5}{(\sqrt{u^2 + v^2 + f^2})^3} e^{-2Z} = 0 \quad (14)$$

Equation (14) is a first order nonlinear partial differential equation namely static H-J equation. Viscosity resolution theory will be used to solve the equation.

3 Fast Sweeping Numerical Scheme to Solve the H-J Equation

The fast sweeping method is based on the Lax-Friedrichs (L-F) monotone numerical Hamiltonians to approximate viscosity solutions of static H-J equations [16-17]. One of the merits of the fast sweeping method is that it can deal with both convex and non-convex functions. Consider the following general static H-J equation

$$\begin{cases} H(z(u, v), p(u, v), q(u, v), u, v) = 0, & (u, v) \in \Omega \setminus \Gamma \\ z(u, v) = g(u, v), & (u, v) \in \Gamma \end{cases} \quad (15)$$

where Ω is the imaging domain to be solved, and Γ is the boundary of Ω . The Hamiltonian H is a nonlinear Lipschitz continuous function, p is $\partial z / \partial u$ and q is

$\partial z / \partial v$. The corresponding vanishing viscosity equation which depending on the parameter $\varepsilon > 0$ is given by

$$\begin{cases} H(z_\varepsilon(u, v), p_\varepsilon(u, v), q_\varepsilon(u, v), u, v) - \varepsilon \Delta z_\varepsilon = 0, & (u, v) \in \Omega \setminus \Gamma \\ z_\varepsilon(u, v) = g(u, v), & (u, v) \in \Gamma \end{cases} \quad (16)$$

The Fast sweeping numerical approximation scheme of Eq.(15) is based on the following lemma.

Lama 1[18]. If the Hamiltonian of Eq.(15) satisfies standard regularity and monotonicity conditions, the solution z_ε of the corresponding vanishing viscosity equations Eq.(16) exist uniquely. Then the solution z_ε will converge to the solution the static H-J equation Eq.(15) such as

$$\lim_{\varepsilon \rightarrow 0} z_\varepsilon(u, v) = z(u, v) \quad (17)$$

Follows give the numerical approximation scheme to solve Eq.(16) by using the corresponding vanishing viscosity solutions of Eq.(17). Let $\{(u_i, v_j), i = 1, 2, \dots, M; j = 1, 2, \dots, N\}$ be the discretized domain of Ω with grid size $(\Delta u, \Delta v)$, and p^\pm and q^\pm be corresponding forward and backward difference approximations of p and q such as

$$\begin{aligned} p_{i,j}^+ &= \frac{Z_{i+1,j} - Z_{i,j}}{\Delta u}, p_{i,j}^- = \frac{Z_{i,j} - Z_{i-1,j}}{\Delta u}, q_{i,j}^+ = \frac{Z_{i,j+1} - Z_{i,j}}{\Delta v}, \\ q_{i,j}^- &= \frac{Z_{i,j} - Z_{i,j-1}}{\Delta v} \end{aligned} \quad (18)$$

And $p = \frac{1}{2}(p^+ + p^-), q = \frac{1}{2}(q^+ + q^-)$, the discretized formulation of the L-F monotone numerical Hamiltonians of Eq.(16) is as

$$H(z(u, v), \frac{p^+ + p^-}{2}, \frac{q^+ + q^-}{2}, u, v) - \varepsilon (\frac{p^+ - p^-}{2} + \frac{q^+ - q^-}{2}) = 0 \quad (19)$$

Then local artificial viscosities factors is used in L-F Hamiltonians Eq.(19), we have

$$H(z(u, v), \frac{p^+ + p^-}{2}, \frac{q^+ + q^-}{2}, u, v) - (\sigma_u \frac{p^+ - p^-}{2} + \sigma_v \frac{q^+ - q^-}{2}) = 0 \quad (20)$$

where σ_u and σ_v are called local artificial viscosities factors which satisfy

$$\sigma_u \geq \max \left| \frac{\partial H}{\partial p} \right|, \quad \sigma_v \geq \max \left| \frac{\partial H}{\partial q} \right|. \quad (21)$$

Because the static H-J equation of SFS problem marked as Eq.(14) is symmetrical re-
 spective with p and q , σ_u and σ_v were set same in our experiments. When Eq.(18)
 are substituted into Eq.(20), we have the following algebra equation

$$H(z(u,v), \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta u}, \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta v}, u, v) - (\sigma_u \frac{Z_{i+1,j} + Z_{i-1,j} - 2Z_{i,j}}{2\Delta u} + \sigma_v \frac{Z_{i,j+1} + Z_{i,j-1} - 2Z_{i,j}}{2\Delta v}) = 0 \tag{22}$$

The concrete formulation of the algebra equation for the SFS problem Eq.(14) is as

$$I(u,v) \sqrt{(f \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta u})^2 + (f \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta v})^2 + (u \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta u} + v \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta v} - (\sigma_u \frac{Z_{i+1,j} + Z_{i-1,j} - 2Z_{i,j}}{2\Delta u} + \sigma_v \frac{Z_{i,j+1} + Z_{i,j-1} - 2Z_{i,j}}{2\Delta v}) - \frac{f^5}{(\sqrt{u^2 + v^2 + f^2})^3} e^{-2Z_{i,j}} \tag{23}$$

From the Eq.(26) we have the following formulation

$$Z_{i,j} = \frac{1}{\frac{\sigma_u + \sigma_v}{\Delta u \Delta v}} ((\sigma_u \frac{Z_{i+1,j} + Z_{i-1,j}}{2\Delta u} + \sigma_v \frac{Z_{i,j+1} + Z_{i,j-1}}{2\Delta v}) + \frac{f^5}{(\sqrt{u^2 + v^2 + f^2})^3} e^{-2Z_{i,j}} - I(u,v) \sqrt{(f \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta u})^2 + (f \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta v})^2 + (u \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta u} + v \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta v} \tag{24}$$

The corresponding iterative computing formula to compute value of $Z_{i,j}$ is

$$Z_{i,j}^{(k+1)} = \frac{1}{\frac{\sigma_u + \sigma_v}{\Delta u \Delta v}} ((\sigma_u \frac{Z_{i+1,j} + Z_{i-1,j}}{2\Delta u} + \sigma_v \frac{Z_{i,j+1} + Z_{i,j-1}}{2\Delta v}) + \frac{f^5}{(\sqrt{u^2 + v^2 + f^2})^3} e^{-2Z_{i,j}} - I(u,v) \sqrt{(f \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta u})^2 + (f \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta v})^2 + (u \frac{Z_{i+1,j} - Z_{i-1,j}}{2\Delta u} + v \frac{Z_{i,j+1} - Z_{i,j-1}}{2\Delta v} \tag{25}$$

where $k \in Z^+$ is the iteration time. The implementation steps of the fast sweeping numerical scheme to solve the perspective SFS H-J equations Eq.(14) are summarized as follows:

1) Initialization: Fix exact values $z_{i,j} = g_{i,j}$ on the boundary Γ , where $g_{i,j}$ is the value at grid point (i, j) in discrete domain. These values are not changed during the iterations. For other pixel points, a large positive value is assigned, the value should be larger than the maximum of the solutions, and these values will be renewed during the process of iterations.

2) Alternating sweeping iteration: At the $k + 1$ th iteration, $Z_{i,j}^{(k+1)}$ are updated using the previous the k th height values according to Eq.(24) at all pixel points except for those with fixed values. And $Z_{i,j}$ is updated only when $Z_{i,j}^{(k+1)}$ is less than the previous value $Z_{i,j}^{(k)}$. This process changes four different directions alternatively as (a) $i = 1 : M ; j = 1 : N$; (b) $i = M : 1 ; j = 1 : N$; (c) $i = 1 : M ; j = N : 1$; (d) $i = M : 1 ; j = N : 1$. And $M \times N$ are the total pixel values of imaging domain Ω .

3) Enforcing computation boundary conditions: On the boundary, a set of conditions should be imposed suggested by [25]. If the boundary of 2-D imaging domain is not rectangle, similar computation is also valid. Linear extrapolation method may be an alternative choice.

4) Stop criteria: If $\|Z_{i,j}^{(k+1)} - Z_{i,j}^{(k)}\| \leq \delta$ where δ is a given positive threshold value, or other stop criteria are satisfied, the computation stops.

Different superscripts of $Z_{i,j}, Z_{i-1,j}, Z_{i+1,j}, Z_{i,j-1}$ and $Z_{i,j+1}$ in right part of Eq.(25) are denoted by using the sweeping directions of Gauss–Seidel iterations to accelerate the convergence of the algorithm. There are four different sweeping directions: From lower left to upper right, from lower right to upper left, from upper left to lower right, and from upper right to lower left. If sweeping is from left-down to right-up (such as $i = 1 : N ; j = M : 1$), then

$$Z_{i-1,j} = Z_{i-1,j}^{(k+1)}, Z_{i,j-1} = Z_{i,j-1}^{(k)} \quad Z_{i+1,j} = Z_{i+1,j}^{(k)}, Z_{i,j+1} = Z_{i,j+1}^{(k+1)} \tag{26}$$

because the newest values are used in the Gauss-Seidel iteration. Other situations are similar when we deal with other three sweeping directions.

4 Experimental Results

Experimental results of both synthetic and benchmark endoscope images are illustrated. The reconstruction results of two real medical endoscope images by using the proposed method and method of reference [15] are performed. All the experiments are realized under the following conditions: hardware CPU Intel Core-2-Duo 1.6GHZ, RAM 1024MB; software Windows XP and Matlab7.0.

4.1 Experimental Result of the Synthetic Image of Surface Function

The synthetic image of vase function Eq.(27) is used. Image is formulated by Eq.(13) using the surface functions, where lighting source was located at $(0, 0, 0)$, and focal length is set as 120. Pixels of images are 200×200 .

$$z(u, v) = f + 0.15 - \sqrt{(0.15 - 0.1v(6v + 1))^2 + (v - 1)^2 + (3v - 2)^2} - u^2, \tag{27}$$

$$(u, v) \in [-0.5, 0.5] \times [0, 1]$$

Figure 1 illustrated experimental results. Original 3-D shapes and synthetic images of vase function are shown as figure 1(a-b). Figure 1(c-d) are reconstructed 3-D shapes and errors with ground true surface by using the proposed method after 50 iterations.

Unit of figures is pixel. Mean error (ME) and root squares mean error (RS) of reconstructed shape with the original shape are 0.0033 and 0.0065 respectively. And CPU time is 261.7530s.

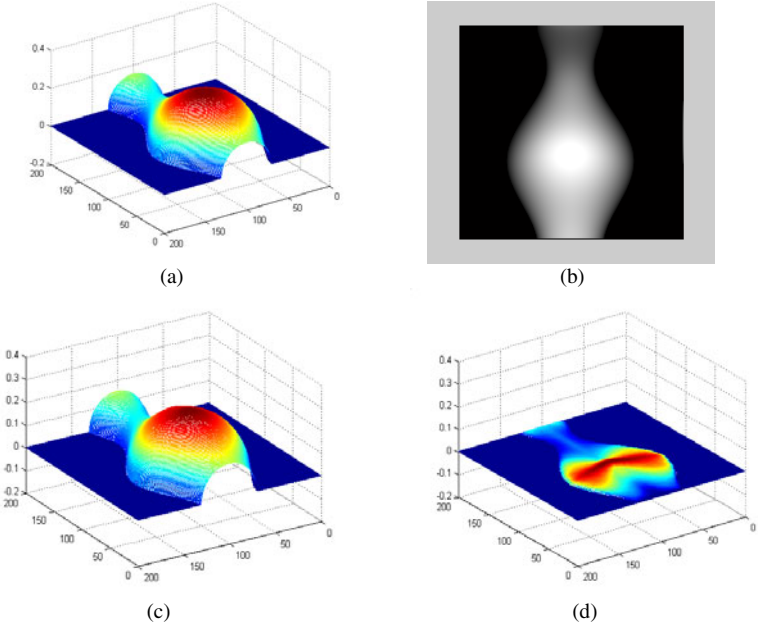
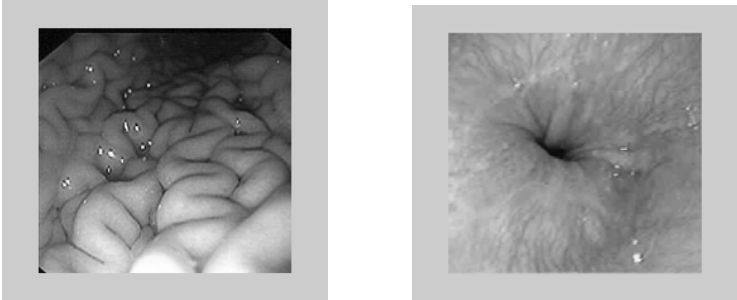


Fig. 1. Reconstruction result of synthetic image of the vase function

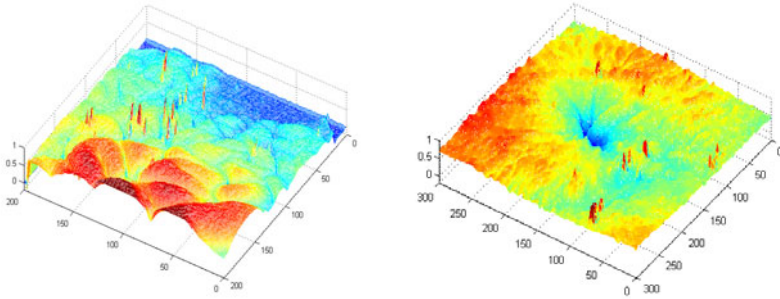
4.2 Experiments with the Benchmark Endoscope Images

We reconstruct the 3-D shapes of two benchmark medical endoscope images by using the proposed imaging model and method of reference [15].

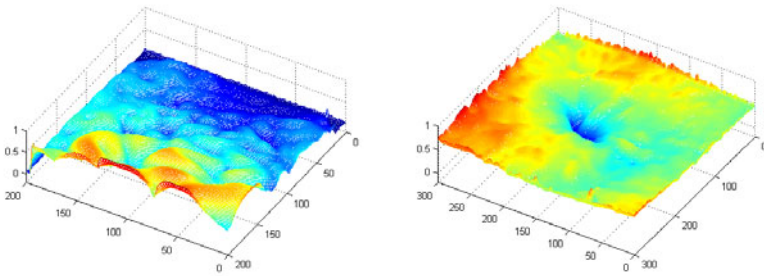


(a) Two medical endoscope images

Fig. 2. Reconstruction results of benchmark endoscope images



(b) Reconstructed shapes by using the proposed method



(c) Reconstructed shapes by using the method of reference [15]

Fig. 2. (continued)

Figure 2(a) shows two endoscope images came from [26] after eight-neighbor mean filtering. Figure 2(b) and (c) show reconstructed shape using the proposed method and the reference [15]. Even though the real 3-D shapes of endoscope images are unknown, we can see that the proposed method is more accurate than the compared method intuitively from figure 2. We notice that the proposed method gives more detail information of shapes form the original images than the reference method.

5 Conclusions

New perspective SFS method has been proposed in this paper. Near-scene point lighting source imaging model under perspective projection is formulated firstly. And the SFS problem is then formulated as viscosity solution of the H-J equation. Following the Lax-Friedrichs fast sweeping numerical scheme based on vanishing viscosity approximation theory is used to solve the H-J equation. At last, experimental results illustrate the performance of the proposed method. As we know, beside the numerical error produced by discrete difference approximation, the error of imaging model is inevitable in SFS methods. The proposed SFS method establishes accurate perspective imaging model with near-scene point lighting source condition.

Acknowledgments. The work is supported by the National Natural Science Foundation of China (61005015), the third National Post-Doctoral Special Foundation of China (201003280), the Post-Doctoral Foundation of China (20100470108), Shanghai University 11th Five-Year Plan 211 Construction Project, and Shanghai Committee of Science and Technology, China (10DZ2273400).

References

1. Horn, B.K.P.: Height and gradient from shading. *Int. J. Computer Vision* 5(1), 37–75 (1990)
2. Zhang, R., Tsai, P.S., Cryer, J.E., Shah, M.: Shape from shading: a survey. *IEEE Trans. PAMI* 21(8), 690–706 (1999)
3. Durou, J.D.: Maurizio Falcone and Manuela Sagona: Numerical methods for shape-from-shading: a new survey with benchmarks. *Computer Vision and Image Understanding* 109(1), 22–43 (2008)
4. Cho, S.Y., Chow, T.W.S.: A new color 3D SFS methodology using neural-based color reflectance models and iterative recursive method. *Neural Computation* 14(11), 2751–2789 (2002)
5. Du, Q.Y., Ben, S.C., Lin, T.: An application of shape from shading. In: 8th International Conference on Control, Automation, Robotics and Vision, pp. 184–189 (2004)
6. Song, L.M., Qu, X.H., Xu, K.X., Lv, L.N.: Novel SFS-NDT in the field of defect detection. *NDT & E International* 38(5), 381–386 (2005)
7. Wu, J., William, A.P.S., Hancock, E.R.: Facial gender classification using shape-from-shading. *Image and Vision Computing* 28(6), 1039–1048 (2010)
8. Bors, A.G., Hancock, E.R., Wilson, R.C.: Terrain analysis using radar shape-from-shading. *IEEE Trans. PAMI* 25(8), 974–992 (2003)
9. Vega, O.E., Yang, Y.H.: Shading logic: A heuristic approach to recover shape from shading. *IEEE Trans. PAMI* 15(6), 592–597 (1993)
10. Lu, J., James, J.L.: Reflectance and shape from images using a collinear light source. *International Journal of Computer Vision* 32(3), 213–240 (1999)
11. Rouy, E., Tourin, A.: A viscosity solutions approach to shape-from-shading. *SIAM Journal of Numerical Analysis* 29(3), 867–884 (1992)
12. Jochen, W., Jörg, K.: Shape from shading using probability functions and belief propagation. *Int J Computer Vision* 84(3), 269–287 (2009)
13. Lin, C.T., Cheng, W.C., Liang, S.F.: Neural-network-based adaptive hybrid-reflectance model for 3-D surface reconstruction. *IEEE Trans. Neural Networks* 16, 1601–1615 (2005)
14. Tankus, A., Sochen, N., Yeshurun, Y.: Shape-from-shading under perspective projection. *International Journal of Computer Vision* 63(1), 21–43 (2005)
15. Yang, L., Han, J.Q.: 3-D shape reconstruction of medical images using perspective shape-from-shading method. *Measurement Science and Technology* 19(6), 502–509 (2008)
16. Prados, E., Faugeras, O.: A generic and provably convergent shape-from-shading method for orthographic and pinhole cameras. *Int. J. Comput. Vis.* 65(1), 97–125 (2005)
17. Breuss, M., Cristiani, E., Durou, J.D., Falcone, M., Oliver, V.: Numerical algorithms for perspective shape from shading. *Kybernetika* 46(2), 207–225 (2010)
18. Bardi, M., Capuzzo-Dolcetta, I.: Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Birkhäuser (1997)

19. Steven, C., Mohamed, A.M.: 3-D ear modeling and recognition from video sequences using shape from shading. *IEEE Trans. Information Forensics and Security* 3(4), 709–718 (2008)
20. Wu, C., Narasimhan, S.G., Jaramaz, B.: A multi-image shape-from-shading framework for near-lighting perspective endoscopes. *Int. J. Comput. Vis.* 86(2-3), 211–228 (2010)
21. Chen, D., Dong, F.: Shape from shading using wavelets and weighted smoothness constraints. *IET Computer Vision* 4(1), 1–11 (2010)
22. Sethian, J.: *Level sets methods and fast marching methods*, 2nd edn. Cambridge U P, Cambridge (1999)
23. Antonio, R.K., Edwin, R.H.: Shape-from-shading using the heat equation. *IEEE Trans. Image Processing* 16(1), 7–21 (2007)
24. Shlizerman, I.K., Basri, R.: 3D face reconstruction from a single image using a single reference face shape. *IEEE Trans. PAMI* 33(2), 394–405 (2011)
25. Zhao, H.: A fast sweeping method for Eikonal equations. *Math. Comp.* 74, 603–627 (2005)
26. <http://www.gastrolab.net/s92enjog.htm>

Mixture of Subspace Learning with Adaptive Dimensionality: A Self-Organizing Approach

Huicheng Zheng

School of Information Science and Technology, Sun Yat-sen University,
135 West Xingang Road, 510275 Guangzhou, China
zhenghch@mail.sysu.edu.cn

Abstract. In computer vision as well as in many other domains, high-dimensional data with potentially low intrinsic dimensions need to be modeled. We propose a subspace mixture model under the self-organizing framework, where each neuron learns an adaptive number of dominant eigenvectors. The overall network constitutes a mixture of ordered local subspaces with the advantage of noise smoothing. Experimental results show that the proposed model can accurately represent high-dimensional visual objects such as handwritten digit images subject to substantial variations, and reveal the intrinsic dimensionality of nonlinearly distributed data.

Keywords: Subspace learning, mixture of subspaces, self-organizing map.

1 Introduction

Subspace learning has been applied in computer vision for invariant representations of visual objects, such as faces, handwritten digits, or general targets [1, 5, 8]. For objects subject to nonlinear transformations, manifold learning may be used. But typical manifold learning approaches lack model generalization. Mixtures of subspaces, on the other hand, have good generalization capacity, which can then be solved by using the expectation-maximization (EM) algorithm.

An alternative solution to the mixture models is biologically inspired. Kohonen proposed the adaptive-subspace self-organizing map (ASSOM) [3], where each unit represents a subspace instead of a reference vector in the SOM [4]. A number of related extensions have been proposed for representing nonlinear manifolds [7, 12]. These methods generally need the dimensionality of local subspaces to be pre-determined in advance, which is often not realistic in practice.

This paper proposes to solve the mixture of subspace learning with adaptive dimensionality under a batch-mode self-organizing framework. At each neuron of the network, the number of eigenvectors for local subspace representation are automatically determined according to the corresponding eigenvalues.

The rest of this paper is organized as follows. Section 2 explains local subspace learning with adaptive dimensionality at each neuron of the network. Section 3 presents the overall mixture model and training algorithm. Experimental results are provided in Sect. 4. Finally, Sect. 5 concludes this paper.

2 Subspace Learning with Adaptive Dimensionality

For high-dimensional data, it is often prohibitive in computational resources to evaluate eigenvectors directly from covariance matrices. Here we adapt the candid covariance-free incremental principal component analysis (CCIPCA) algorithm proposed by Weng et al. [9] to compute an adaptive number of eigenvectors.

2.1 The Candid Covariance-Free Incremental PCA

The basic idea of CCIPCA [9] can be summarized as follows. Let \mathbf{x} be a d -dimensional random vector with zero mean, $\mathbf{x}(1), \mathbf{x}(2), \dots$ be a sequence of samples from \mathbf{x} , $\Sigma = E[\mathbf{x}\mathbf{x}^T]$ be the covariance matrix. Then an eigenvector \mathbf{e} of Σ satisfies $\lambda\mathbf{e} = \Sigma\mathbf{e}$, where λ is the corresponding eigenvalue. Replacing the unknown Σ with its estimate, and \mathbf{e} by its incremental estimates, we obtain the following expression:

$$\lambda(n)\mathbf{e}(n) = \frac{1}{n} \sum_{i=1}^n \mathbf{x}(i)\mathbf{x}^T(i)\mathbf{e}(i-1). \tag{1}$$

Equation (1) can be written in a recursive form for algorithmic computation:

$$\begin{aligned} \mathbf{v}(n) &= \frac{n-1}{n}\lambda(n-1)\mathbf{e}(n-1) + \frac{1}{n}\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{e}(n-1), \\ \lambda(n) &= \|\mathbf{v}(n)\|, \mathbf{e}(n) = \mathbf{v}(n)/\lambda(n), \quad n = 2, 3, \dots, \end{aligned} \tag{2}$$

where $\lambda(1) = \|\mathbf{x}(1)\|$, $\mathbf{e}(1) = \mathbf{x}(1)/\lambda(1)$. Convergence of this algorithm has been proved in [11]. In short, $\mathbf{e}(n)$ tends to the unit-length principal eigenvector, and $\lambda(n)$ to the corresponding eigenvalue, when $n \rightarrow \infty$. If we define $\mathbf{u}(i) = \mathbf{x}(i)\mathbf{x}^T(i)\mathbf{e}(i-1)$ in (1), then $\mathbf{v}(n)$ is an efficient estimate under the Gaussian assumption. That is, the estimate tends to converge most quickly with the least error variance. Since most realistic distributions are close to Gaussian, the efficiency of the estimate is at least approximately true in most situations. Subsequent higher-order eigenvectors can be computed in the complementary space by subtracting from the data their projections on the preceding dominant eigenvectors.

2.2 Subspace Learning with Adaptive Dimensionality

The CCIPCA in its original form can be used to compute a pre-determined number of dominant eigenvectors. To automatically determine the dimensionality of the embedded principal subspace, we may take advantage of the computed eigenvalues. For example, the ratio of the current eigenvalue to the leading eigenvalue can be compared to a threshold T_s to decide if more eigenvectors should be computed.

The task of determining the subspace dimensionality is facilitated under batch-mode learning. In this case, the mean of sample vectors can be accurately estimated, without the nuisance of drifting in the online learning. Furthermore, each subsequent eigenvector can be estimated based on the *convergent* preceding eigenvectors, which is expected to be more accurate and converge faster than the original CCIPCA. In fact, in the original CCIPCA, eigenvectors estimated before convergence could break mutual orthogonality.

The batch-mode algorithm for computing the principal subspace while automatically determining the embedded dimensionality can be summarized as follows.

1. Compute the leading eigenvector \mathbf{e}_1 and the corresponding eigenvalue λ_1 by using the CCIPCA algorithm. Subtract from the sample vectors their projections on \mathbf{e}_1 , i.e. for each sample vector $\mathbf{x}_1(n) = \mathbf{x}(n)$, compute $\mathbf{x}_2(n) = \mathbf{x}_1(n) - \mathbf{x}_1^T(n)\mathbf{e}_1(n)\mathbf{e}_1(n)$.
2. For $i = 2, \dots$, do
 - (a) Compute the i -th eigenvector \mathbf{e}_i and the corresponding eigenvalue λ_i by using all the current sample vectors $\mathbf{x}_i(n)$.
 - (b) If $\lambda_i/\lambda_1 > T_s$, subtract from $\mathbf{x}_i(n)$ their projections on \mathbf{e}_i , i.e. compute $\mathbf{x}_{i+1}(n) = \mathbf{x}_i(n) - \mathbf{x}_i^T(n)\mathbf{e}_i(n)\mathbf{e}_i(n)$ for all $\mathbf{x}_i(n)$.
 - (c) Else break.

As for the computational complexity, the batch-mode algorithm does not introduce extra computational cost compared to the original CCIPCA algorithm.

3 Mixture Learning with a Self-Organizing Approach

The mixture model is motivated by the self-organizing map (SOM) [4], which consists of a grid of processing units each representing a local prototype of input patterns. In the learning phase, for each input vector, a winner among the neurons is determined according to its similarity to the input vector. The winner and its neighbors (as defined by a neighborhood function) are then updated by the input vector. The generated map is automatically ordered at the convergent stage.

We use a batch-mode self-organizing approach in this work. The batch mode contains no learning-rate parameter as opposed to the original algorithm, and therefore has no convergence problems and yields stabler learning behavior [2]. It works by partitioning the training data first, similar to the K -means algorithm. Then each neuron is updated by the training data in the partitions of the current neuron and the neighboring neurons. It has been shown that the batch-mode SOM shows better robustness than the K -means algorithm [2]. The neighborhood set reduces with learning iterations, which enforces ordering of neurons in the early stage and accurate local learning in the final stage. The batch-mode algorithm generally converges in a few iterations.

To partition the data space, we model the principal subspace at each neuron with a Gaussian distribution. The similarity is then defined as a likelihood

function. By using a likelihood measure instead of the full Bayesian posterior, we are in effect implementing a uniform prior on the neurons, which leads to a smoothing effect that improves the robustness.

Let \mathbf{m} be the mean vector, $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d$ be the unit eigenvectors, and $\lambda_1, \lambda_2, \dots, \lambda_d$ be the corresponding eigenvalues at an arbitrary neuron in the network, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$. The orthogonal matrix $\mathbf{U} = [\mathbf{e}_1 \mathbf{e}_2 \dots \mathbf{e}_d]$, $\mathbf{A} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_d]$. Then the local covariance matrix $\mathbf{\Sigma} = \mathbf{U} \mathbf{A} \mathbf{U}^T$. Since only the first k dominant eigenvectors are evaluated, we replace $\lambda_{k+1}, \dots, \lambda_d$ by the average value $\lambda_0 = \sum_{i=k+1}^d \lambda_i / (d - k)$, which involves the unknown eigenvalues. To evaluate λ_0 in practice, we consider the residual vector $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$, where $\hat{\mathbf{x}} = \mathbf{m} + \sum_{i=1}^k ((\mathbf{x} - \mathbf{m})^T \mathbf{e}_i) \mathbf{e}_i$ is the projection of \mathbf{x} on the local subspace spanned by the first k eigenvectors. Then $\sum_{i=k+1}^d \lambda_i = E[\tilde{\mathbf{x}}^T \tilde{\mathbf{x}}]$. Therefore, we have the estimate:

$$\lambda_0 = \frac{\sum_{n=1}^N \tilde{\mathbf{x}}^T(n) \tilde{\mathbf{x}}(n)}{N(d - k)}, \quad (3)$$

where $\tilde{\mathbf{x}}(n)$ are the residual vectors corresponding to the training vectors $\mathbf{x}(n)$, $n = 1, 2, \dots, N$ in the partitions of the current neuron neighborhood set.

Now we define $\mathbf{\Sigma} = \mathbf{U} \mathbf{A}' \mathbf{U}^T$, where $\mathbf{A}' = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_k, \lambda_0, \dots, \lambda_0]$. Plugging it into the Gaussian density function, we obtain the formula for computing the log-likelihood of an input vector \mathbf{x} given a local k -dimensional principal subspace:

$$L(\mathbf{x}) = -\frac{1}{2} \left(d \ln 2\pi \lambda_0 - \sum_{i=1}^k \ln \frac{\lambda_i}{\lambda_0} - \sum_{i=1}^k \frac{((\mathbf{x} - \mathbf{m})^T \mathbf{e}_i)^2}{\lambda_i} - \frac{\tilde{\mathbf{x}}^T \tilde{\mathbf{x}}}{\lambda_0} \right). \quad (4)$$

A uniform eigenvalue λ_0 smoothes noise in the less reliable ‘‘residual’’ dimensions and helps to improve the robustness of the algorithm.

The overall algorithm works in iterations and can be summarized as follows.

1. Initialize the mean vector of each neuron with a training vector.
2. Partition the training vectors according to their distances to the mean vectors.
3. At each neuron, compute the mean vector, then the dominant eigenvectors one by one based on the training vectors in the partitions of the current neuron and its neighbors, until the stop criterion is satisfied, as explained in Sect. 2.2
4. Partition the training vectors according to their similarity to the neurons with the likelihood measure.
5. Update the neighborhood set. At each neuron, compute the mean vector, then the dominant eigenvectors one by one based on training vectors in the partitions of the current neighborhood set, until the stop criterion is satisfied.
6. Repeat 4–5 a few times.

It could happen accidentally that some partitions contain insufficient training vectors. In this case, the neighborhood set can be temporarily enlarged to include more data. Thus noise in the learning process can be smoothed.

4 Experimental Results

The proposed mixture of subspaces can be used to model various visual objects, e.g. faces or handwritten digits subject to nonlinear transformations. The experimental data set used here is the publicly available MNIST database [6], which contains images of handwritten digits “0”–“9”. There are 60,000 training images and 10,000 test images in the database, which have been size-normalized and centered in a 28×28 pixel grid. Pixel values are in grayscale due to the antialiasing technique used in normalization. Figure 1 shows some examples of one digit from the MNIST database, which apparently subject to substantial variations.



Fig. 1. Examples of digit “8” from the MNIST database

A network is trained for each digit. For each digit image, the mean value of its pixels has been subtracted from each pixel. Then the image as a vector is normalized to have unit length. For the first half of training iterations, the radius of the neighborhood set decreases linearly with the iterations, starting with a value that covers the full network. For the remaining iterations, the neighborhood set only contains the current neuron to enforce an accurate local learning. In our experiments, the total number of training iterations is set to be 10. More iterations did not improve the performance significantly.

Figure 2 shows the network of size 6×6 trained by using the images of digit “8”. The ratio T_s is 0.02, i.e. the network stops learning more eigenvectors whenever one of the neuron reports a ratio between the current and the leading eigenvalues lower than T_s . As we can see in Fig. 2, 6-dimensional local subspaces have been generated. The mean vectors and eigenvectors are ordered spatially, each one being adapted to a local partition showing a specific “style”.

Figure 3 shows the mean values and standard deviations of the first 15 eigenvalues among the neurons. We can see that the eigenvalues are automatically ordered by their magnitudes. The variance in the remaining $28 \times 28 - 15 = 769$ dimensions is only 0.144, or 1.87×10^{-4} averaged on each dimension.

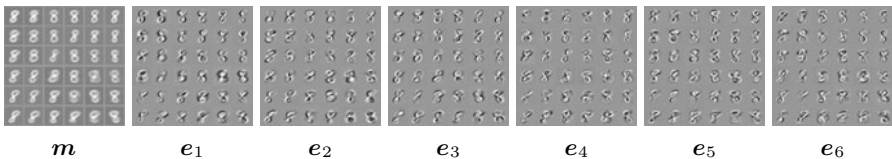


Fig. 2. Mixture of 6-dimensional subspaces learned from digit “8” in the MNIST training set. m and $e_i, i = 1, 2, \dots, 6$ denote, respectively, the mean vectors and the corresponding eigenvectors that span the local subspaces. The eigenvectors are shown in the descending order of corresponding eigenvalues.

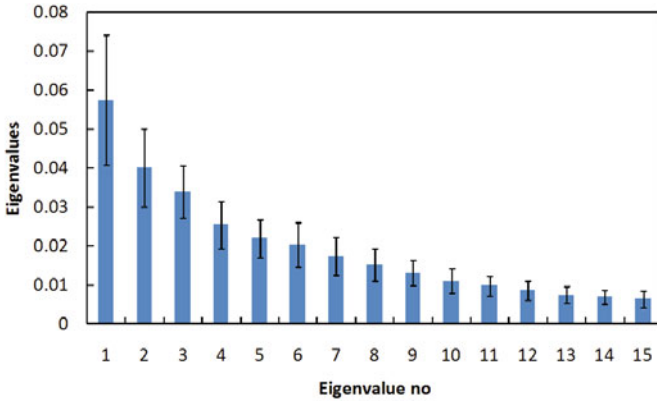


Fig. 3. Distributions of the first 15 eigenvalues learned from digit “8”

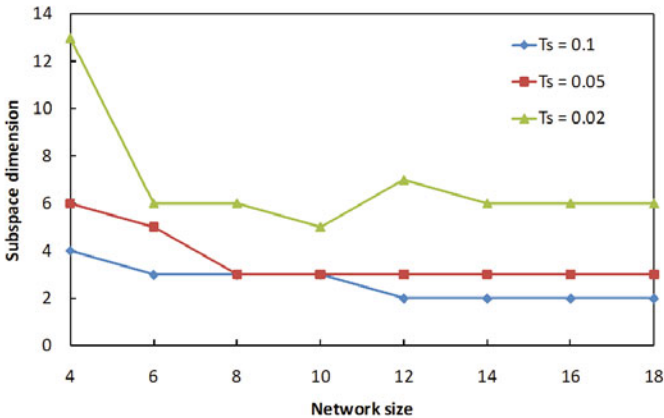


Fig. 4. Subspace dimensionality varies with the network size under various T_s values. The horizontal axis represents the number of rows (or columns) of neurons in the square networks. The vertical axis represents the dimensionality of local subspaces.

The dimensionality of local subspaces increases with decreasing T_s , as demonstrated in Fig. 4. On the other hand, as a general tendency, the subspace dimensionality decreases with the increasing network size under a fixed T_s . The reason could be that, with increasing number of partitions, local data structures become simpler, and can be well described by less dimensions. However, after some certain point, the subspace dimensionality does not seem to decrease further with the increasing network size. It seems that there may exist some stable fundamental subspace dimensionality under the fixed threshold T_s .

For a test digit image, a reconstructed image can be computed from each network. By comparing the input image with the “memorized” image in the individual networks, we may identify the input digit. Let \hat{x}_q be the projection

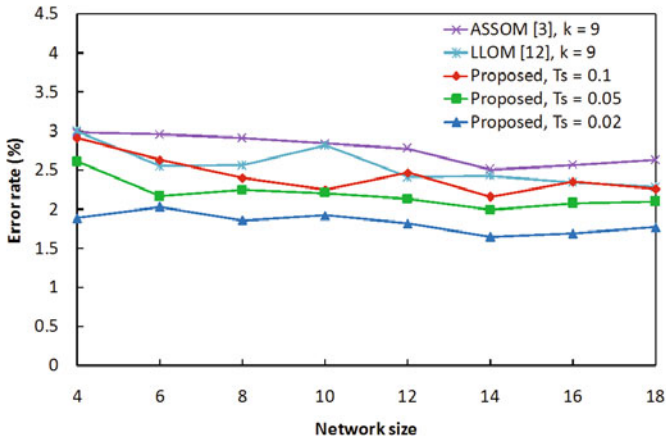


Fig. 5. Error rates of the proposed algorithm compared to those in the literature on the MNIST test set

of the input image \mathbf{x} on the local subspace of the q -th neuron, where $q \in Q$ with Q being the set of neurons in a network. The reconstructed image from the network is computed as $\hat{\mathbf{x}}_Q = \sum_{q \in Q} a_q \hat{\mathbf{x}}_q / \sum_{q \in Q} a_q$, where the weight a_q is inversely related to the distance between \mathbf{x} and its projection $\hat{\mathbf{x}}_q$, as in [10]. A Gaussian function of the Euclidean distance is chosen for a_q .

Recognition results on the test set have been summarized in Fig. 5, which are promising compared to other approaches in the literature. Generally, the lower the threshold T_s is, i.e. the more information we preserve, the lower error rate we may obtain. The error rates are less related to the size of the networks.

5 Conclusions

This paper proposes a mixture of subspaces model realized under the self-organizing framework. Its main property is that the intrinsic dimensionality of local subspaces can be automatically determined from the training data. The self-organizing framework brings the advantages of ordered low-dimensional (often 2D) visualization and noise smoothing, which also contributes to the robustness of the algorithm. The proposed model can be used as an accurate representation of objects under complex nonlinear transformations, as demonstrated in the experiments.

Acknowledgments. This work was supported by National Natural Science Foundation of China (61172141), Specialized Research Fund for the Doctoral Program of Higher Education (20080558-1005), Fundamental Research Funds for the Central Universities (09lgpy52), Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, and Innovative Research Fund for Outstanding Young Scholars of Guangdong Higher Education.

References

1. Cai, D., He, X., Hu, Y., Han, J., Huang, T.: Learning a spatially smooth subspace for face recognition. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–7 (2007)
2. Kohonen, T.: Things you haven't heard about the self-organizing map. In: Proceedings of IEEE International Conference on Neural Networks, pp. 1147–1156 (1993)
3. Kohonen, T.: The adaptive-subspace SOM (ASSOM) and its use for the implementation of invariant feature detection. In: Proceedings of International Conference on Artificial Neural Networks, pp. 3–10 (1995)
4. Kohonen, T.: Self-Organizing Maps, 3rd edn. Springer, Heidelberg (2001)
5. Laaksonen, J., Oja, E.: Subspace dimension selection and averaged learning subspace method in handwritten digit classification. In: Proceedings of International Conference on Artificial Neural Networks, pp. 227–232 (1996)
6. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 2278–2324 (1998)
7. Liu, Z.Q.: Adaptive subspace self-organizing map and its application in face recognition. International Journal of Image and Graphics 2, 519–540 (2002)
8. Wang, D., Zhang, C., Hao, P.: Heavy-tailed model for visual tracking via robust subspace learning. In: Proceedings of Asian Conference on Computer Vision, pp. 172–181 (2010)
9. Weng, J., Zhang, Y., Hwang, W.S.: Candid covariance-free incremental principal component analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 25, 1034–1040 (2003)
10. Zhang, B., Fu, M., Yan, H., Jabri, M.: Handwritten digit recognition by adaptive-subspace self-organizing map (ASSOM). IEEE Transactions on Neural Networks 10, 939–945 (1999)
11. Zhang, Y., Weng, J.: Convergence analysis of complementary candid incremental principal component analysis. Technical Report MSU-CSE-01-23, Dept. of Computer Science and Eng., Michigan State Univ., East Lansing (August 2001)
12. Zheng, H., Shen, W., Dai, Q., Hu, S., Lu, Z.M.: Learning nonlinear manifolds based on mixtures of localized linear manifolds under a self-organizing framework. Neurocomputing 72, 3318–3330 (2009)

Dialect Identification Based on S-Transform and Singular Value Decomposition

Yan He and Fengqin Yu

School of Internet of things engineering, Jiangnan University, Wuxi, China
heyang_19871005@163.com, yufq@jiangnan.edu.cn

Abstract. A speech signal being non stationary, the S-transform which combines short-time Fourier analysis and wavelet analysis is an effective tool for analyzing a non stationary signal. Because of the high dimension of time-frequency representations, the singular value decomposition is used to extract the features vectors. In the simulation experiment, firstly, STFT, WT and the S-transform are used to analyze speech signals respectively, and the results show that the time-frequency distribution using the S-transform performs best. Then the right and left singular vectors are extracted using SVD from the time-frequency distributions of Mandarin, Shanghainese, Cantonese and Hokkien as the input feature vectors for SVM to recognize. And the simulation experiment results show that the recognition rate using the S-transform can reach up to 82.5%, higher than using STFT and WT.

Keywords: dialect identification, time-frequency analysis, S-transform, singular value decomposition.

1 Introduction

Dialect identification is the technology of identifying the dialect from a short utterance of the unknown speech signal. And it is extremely important to extract the most important features of speech signals [1]. The dialect speech signal being a non stationary signal, the S-transform time-frequency representation is an appropriate method to extract the important features from non stationary signals. The S-transform as a hybrid of short-time Fourier analysis and wavelet analysis employs a variable window length but preserves the phase information by using the Fourier kernel in the signal decomposition [2]. And the S-transform has already found applications in many fields such as geophysics, power quality analysis, and signal processing for mechanical systems [3].

However, due to the high dimension of time-frequency representations, the success relies upon an appropriate form of dimensionality reduction. And the singular value decomposition (SVD) provides an effective means of concentrating that information which is important, and discarding that which is irrelevant. Singular vectors (SVs) of the time-frequency distribution are the span bases of the matrix, and their importance in the composition of the matrix is reflected by the related singular values [4].

In this paper, the short time Fourier transform (STFT), the wavelet transform(WT) and S-transform are used to analyze one speech signal separately, and the results

show that the S-transform time-frequency distribution performs best. Then because of the high dimension of time-frequency representations, the singular value decomposition (SVD) is employed to extract the feature vector from the time-frequency distributions of Mandarin, Shanghainese, Cantonese and Hokkien to recognize with SVM, and the simulation experiment results show that the recognition rate using the S-transform can reach up to 82.5% higher than that of using STFT or WT, but its increase is based on the cost of computing time.

2 The Basic Theory

2.1 S-Transform

The time-frequency representation is a powerful tool for extracting features of a non stationary signal. And the S-transform is a new approach of the time-frequency representation proposed by Stockwell et al. in [5], which combines the short time Fourier transform (STFT) and wavelet transform (WT). The S-transform uses frequency-dependent Gaussian window function with a multi-resolution, to overcome the lack of the fixed resolution of the STFT. While the S-transform contains phase factor, which is that the WT are not available.

S-transform can be written as operations on the Fourier spectrum $X(f)$ of $x(t)$

$$S(\tau, f) = \int_{-\infty}^{+\infty} X(\alpha + f) e^{\frac{2\pi^2\alpha^2}{f^2}} e^{j2\pi\alpha\tau} d\alpha \quad (f \neq 0). \tag{1}$$

Let T donate a time sampling interval and t donate the sampling time, then the total number of the sampling points is $N = t/T$. So the S-transform of a discrete time series $h[kT]$ is given by (letting $f \rightarrow n/NT$ and $\tau \rightarrow jT$)

$$S[kT, \frac{n}{NT}] = \sum_{m=0}^{N-1} X[\frac{m+n}{NT}] e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{j2\pi mk}{N}} \quad (k, m = 0, 1, \dots, N-1; n = 1, 2, \dots, N-1) \tag{2}$$

And for the $n = 0$ voice, it is equal to the constant defined as

$$S[jT, 0] = \frac{1}{N} \sum_{m=0}^{N-1} x(\frac{m}{NT}). \tag{3}$$

2.1.3 Singular Value Decomposition

SVD has been a valuable tool in signal processing and statistical data analysis [6]. Suppose that $A_{m \times n}$ ($m > n$) is a real matrix, and $\text{rank}(A) = k$, then A could be expressed by m -order orthogonal matrix U and n -order orthogonal matrix V

$$A = UDV^T, U^*U^T = I, V^*V^T = I. \tag{4}$$

Where $D_{m \times n} = \begin{bmatrix} \Sigma_{k \times k} & 0 \\ 0 & 0 \end{bmatrix}$ and $\Sigma_{k \times k} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k)$. Then $\sigma_i = \sqrt{\lambda_i}$

($i = 1, 2, \dots, k, \dots, n$) are the singular values of A, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$ are all

the non-zero eigenvalues of $A^* A^T$ or $A^T * A$, $\lambda_{k+1} \geq \lambda_{k+2} \geq \dots \geq \lambda_n > 0$ are $n-k$ zero eigenvalues of $A^T * A$ and $u_i, v_i (i=1, 2, \dots, k)$ is the eigenvector of $A^* A^T$ or $A^T * A$, respectively. And where $U_{m \times m} = (u_1, u_2, \dots, u_{k+1}, \dots, u_m)$ and $V_{n \times n} = (v_1, v_2, \dots, v_{k+1}, \dots, v_n)$ are orthonormal matrices. The columns of the orthonormal matrices U and V are called the left and right singular vectors (SVs), respectively. An important property of U and V is that they are mutually orthogonal. The singular values (σ_{ii}) represent the importance of individual SVs in the composition of the matrix. In other words, SVs corresponding to the larger singular values have more information about the structure of patterns embedded in the matrix than the other SVs.

However, singular values do not carry significant information embedded in the matrix. In other words, they are not suitable features for classification purposes. To find the characteristics of a signal in the time-frequency domain using the SVD technique, we propose to use the right and left SVs corresponding to the largest singular values. The reason is that the right and left SVs contain the time and frequency domain information of the signal respectively. In addition, SVs related to the largest singular values have more information about the structure of the signal. Consequently, if the structures of signals are different for dissimilar classes, using SVs related to the largest singular values is more suitable for classification. However, SVs corresponding to the lowest singular values would be more appropriate if the structures of different classes are similar to each other.

3 The Implementation

Previous researches have shown that the S-transform with a good ability of time-frequency analysis can provide the distribution information of the signal amplitude changing with time and frequency, while the left and right SV which are extracted from the S-transform time-frequency matrix with SVD can reduce the dimension of feature space and reflect the characteristics of a signal in the time-frequency domain. In the analysis of signals in the time-frequency domain using SVD, the type of time-frequency distribution is important. Indeed, it is desirable that the time-frequency representation is linear and has high resolution, which is the case of the S-transform. To find the characteristics of a signal in the time-frequency domain using the SVD technique, we propose to use the left and right SVs of the time-frequency distribution of the signals. The reason is that the right and left SVs contain the time and frequency domain information of the signal respectively. And the steps of the feature vector extraction with the S-transform and SVD are as follows:

- 1) The signals of four dialects are read and pre-processed including sampling, pre-emphasis, adding hamming windows and so on. In this paper, $1-0.9375z^{-1}$ is employed as the pre-emphasis filter.

- 2) Calculate the Fast Fourier Transform (FFT) $H[\frac{m}{nT}]$ of the every speech signal $x(i)$ and then augment $H[\frac{m}{nT}]$ into $H[\frac{m+n}{nT}]$.
- 3) Calculate the Fast Fourier Transform $W(m, n) = e^{-\frac{2\pi^2 m^2}{n^2}}$ of $w(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$.
- 4) Calculate $H[\frac{m+n}{nT}]W(m, n)$ by frequency sampling points.
- 5) Calculate the Fast Fourier Inverse Transform of $H[\frac{m+n}{nT}]W(m, n)$ and the result is the S-transform matrix $S_i[m, n]$, where i is the number of the speech signals.
- 6) Let $A_i = S_i[m, n]$ and decompose every A_i into $A_i = U_i D_i V_i^T$, Where $D_i = \begin{bmatrix} \Sigma_{k \times k} & 0 \\ 0 & 0 \end{bmatrix}$, $\Sigma_{k \times k} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k)$, $U_{m \times m} = (u_1, u_2, \dots, u_{k+1}, \dots, u_m)$ and $V_{n \times n} = (v_1, v_2, \dots, v_{k+1}, \dots, v_n)$. In this paper, the left and right SVs are used as the input features for SVM to recognize.

4 The Experiment Results

4.1 Time-Frequency Distributions

In the experiment, the speech signal “wo” is studied. To demonstrate the S-transform would be useful as an effective tool to analyze a non stationary signal, STFT, WT and S-transform is adopted to analyze this signal separately. And Fig.1. is the three time-frequency distributions.

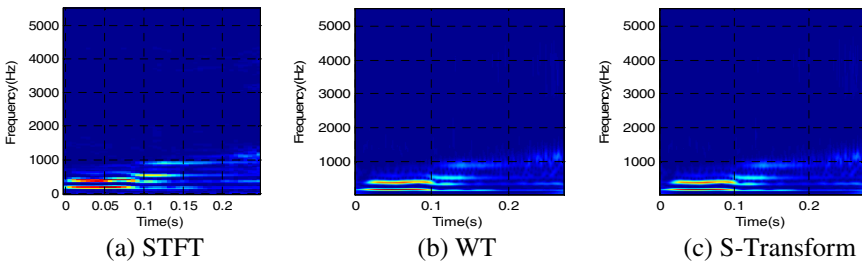
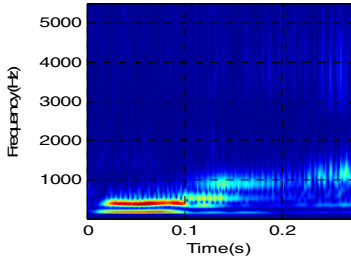


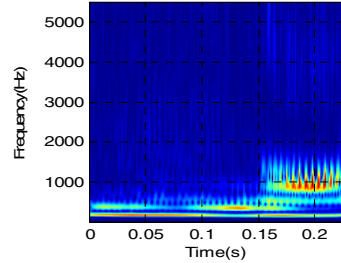
Fig. 1. Time-Frequency Distributions. Seen from Fig.1.(a), STFT with a fixed window length has a contradiction in time resolution and spectral resolution, so the time-frequency distribution generally reflects the distribution of concentration, but is very vague. In Fig.1.(b), WT with multi-resolution provides the speech signal with an adaptive window, the performance of the time-frequency distribution significantly improves compared to that of STFT. Fig.1.(c) shows that the time-frequency distribution has higher time-frequency resolution and better energy aggregation.

4.2 The S-Transform Time-Frequency Distributions

In this experiment, the Chinese word ‘wo’ is studied and is read with Mandarin, Shanghainese, Cantonese and Hokkien separately. Then S-transform is used to analyze the time-frequency distributions. And Fig.2. is the four S-transform time-frequency distributions.

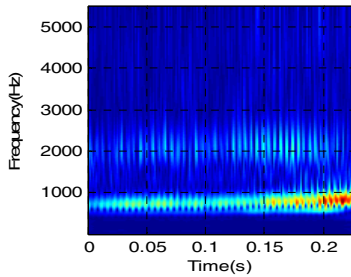


(a) The S-Transform of Mandarin

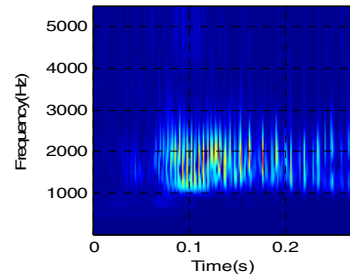


(b) The S-Transform of Shanghai-

nese



(c) The S-Transform of Cantonese



(d) The S-Transform of Hokkien

Fig. 2. S-Transform Time-Frequency Distributions. This is shown that for the different pronunciations of the same word, the S-transform time-frequency distribution of every dialect is different. So it is reasonable to believe that the use of the S-transform for the dialect identification is possible.

4.3 The Recognition Performances

The database of dialect identification comes from the web. The speakers of each dialect are one man and one woman, and in all there are 240 utterances in the database, in which 80 utterances are for testing. Mandarin, Shanghainese, Cantonese and Minnanese is used as the research objects, and speech signals are sampled with 11025Hz sampling frequency and 16-bit sampling precision. In this paper, the S-transform is used to analyze the speech signals. And then SVD is adopted to extract the left and right SVs from the time-frequency matrix to recognize with SVM. In order to compare, the STFT and the WT are used to analyze the speech signals separately. And Table 1 is the comparison of the recognition performance.

Table 1. The Recognition Performances

The method	The recognition rate	The running time
STFT	75%	296.5s
WT	80%	587.34s
S-transform	82.5%	776.734s

The comparison of the recognition performances is shown in table 1. The average recognition rate using STFT is lower than that using WT, while the average recognition rate using WT is lower than that using the S-transform. That is because WT with multi-resolution provides the speech signal with an adaptive window, the performance of the time-frequency distribution significantly improves compared to that of STFT, and the S-transform has high resolution and energy aggregation. However, the running time using the S-transform is the most of all.

Acknowledgments. I would like to express my gratitude to all those who helped me during the writing of this thesis. First of all, I give my most sincere thanks to the National Natural Science Foundation of China under Grant No.61075008. Then I would like to extend my sincere gratitude to professor Fengqin Yu for her instructive advice and useful suggestions on my thesis.

References

1. Gu, M.L., Xia, Y.G., Yang, Y.M.: Semi-supervised learning based Chinese dialect identification. In: Proc. 9th International Conference on Signal Processing, Beijing, pp. 1608–1611 (2008)
2. Ercin, S., Stankovic, L.J., Milos, D., Jin, J.: Instantaneous frequency estimation using the S-transform. *IEEE Signal Processing Letter* 15, 309–312 (2008)
3. Sergi, V., Carine, S., Martin, S., Juan Jose, D., Antoni, M.: The S-transform from a wavelet point of view. *IEEE Transactions on Signal Processing* 56(7), 2771–2780 (2008)
4. Said, A., Anne, H., Jean, P.L.: S-transform time-frequency feature extraction of laser Doppler flowmetry signal using SVD decomposition. In: Proc. ICASSP, pp. 157–160 (2006)
5. Stockwell, R.G., Mansinha, L., Lowe, R.P.: Localization of the complex spectrum: the S-transform. *IEEE transactions on signal processin* 44(4), 998–1001 (1996)
6. Hamid, H., Mostefa, M., Boualem, B.: Time-frequency feature extraction of newborn EEG seizure using SVD- based techniques. *EURASIP Journal on Applied Signal Processing* 16, 2544–2554 (2004)

Local Structure Recognition of Point Cloud Using Sparse Representation

Pei Luo, Zhuangzhi Wu, and Teng Ma

Department of Computer Science & Engineering, Beihang University,
Beijing 100191, China

areslp@cse.buaa.edu.cn, zzwu@buaa.edu.cn, mateng@cse.buaa.edu.cn

Abstract. The local structure of point cloud is a key problem in point based geometry processing. In this paper, we propose a dictionary learning based method to extract the local structure. The core idea is: As point cloud can be seen as a linear model in local view, we use the union of multi-subspace to approximate it. An overcomplete dictionary D is used to represent the bases of these subspaces. First, we calculate the neighborhood N of each point by k -NN and build EMST on it, marked as T . Then, each edge in T is used to construct a training set. Most of the samples in training set indicate the trend of the point set. At last, we solve the sparse matrix factorization problem recursively to update D until D stops changing. We present 2D/3D experimental results to show that this method can handle manifold/non-manifold structures.

Keywords: dictionary learning, local structure, point cloud, sparse representation.

1 Introduction

Point cloud has emerged in recent years as a versatile representation for geometric models in computer graphics. The surface of a 3D object is described by a set of sample points without further topological information such as triangle mesh connectivity or a parameterization. The local structure of point cloud is a key problem of point based 3D geometry processing. It is strongly related to normal estimation[1], surface reconstruction[2], model denoising[3] problems and so on. It is also a common problem occurs in computer vision, image processing, pattern recognition, and systems identification if the ambient space is n -dimensional.

Traditional methods for structure estimation, such as PCA(Principal Component Analysis), ICA(Independent Components Analysis), all have natural defects of their own. The basis functions of PCA are restricted to be orthogonal, while ICA demands that the number of basis vectors is equal to the number of inputs. For example, as shown in Fig 1, PCA and ICA may be influenced by points in other subspace (yellow and pink points). But with local structure information, we can estimate the normal only with red points. To overcome the defects of ICA, Lewicki and Sejnowski[4] presented an algorithm for learning an overcomplete basis by viewing it as probabilistic model of the input data. More recently, Vidal and Ma[5] present an algebro-geometric solution to this kind of problem.

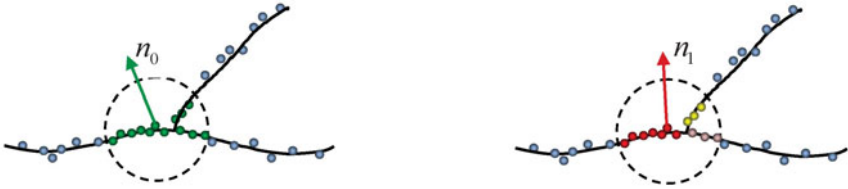


Fig. 1. Defects of traditional methods. This is an example of normal estimation in 2D. The first column shows the normal estimated by PCA or ICA. The second column shows the result by local structure recognition. n_0 is the normal estimated by PCA/ICA, n_1 is the normal estimated by local structure method. PCA will fail when data does not have gaussian distribution, while ICA can not handle three-armed structure in 2D.

In contrast, we will propose an iterative method in sparse view. Using sparse representation, our approach can automatically determine the number of subspaces, which is similar with the k-chosen problem in data clustering. Besides, the method will converge quickly, as shown in our experiments, the algorithm stops after a few iterations.

This paper is organized as follows. Section 2 provides the background on Dictionary Learning, while Section 3 summarizes existing schemes for local structure analysis. Section 4 develops our proposed local structure extract algorithms, and Section 5 presents our experimental results. Section 6 closes with conclusions.

2 Background

In recent years there is a growing interest in the study of sparse representation for signals. Using an overcomplete dictionary that contains prototype signal atoms, signals are described by sparse linear combinations of these atoms.

We refer to the columns of the full row-rank $m \times n$ matrix $D = [d_1, d_2, \dots, d_n]$ as a dictionary. Dictionary learning[6–8] is concentrating on finding a dictionary which can decompose signals sparsely. Assume that a signal $x \in R^m$ can be represented as a sparse linear combination of columns of D, we can write this as $x = D\alpha$, α displays the coefficients of the signal x . So, α is the solution of P(1).

$$\min_{\alpha} \|\alpha\|_0 \quad s.t. \quad x = D\alpha \quad . \tag{1}$$

There are lots of algorithms to learn a dictionary, for example, MOD (Methods of Optimal Directions) by Engan[10], K-SVD by Aharon and Elad[9], and recently, Online dictionary learning methods by Mairal[6].

3 Proposed Method

Point cloud can be seen as a special type of signal, and almost all point clouds have well-marked geometry structure. The idea of sparse representation of point cloud is reasonably straightforward.

Local part of point cloud can be well approximated by linear subspaces. Denote the local part of point cloud as P , it's the union of subspaces $\{S_1, S_2, \dots, S_k\}$. For every point $p \in P, p \in S_i (i \in \{1, \dots, k\})$.

We will formulate the local structure analysis problem as an iterative dictionary learning problem. Firstly, all bases of subspaces will be extracted. Then, these bases will be clustered into different groups by the subspace identification step.

Assume the columns of D is the overcomplete bases of P , and d_1, \dots, d_{s_1} is the bases of subspace $S_1, d_{s_1+1}, \dots, d_{s_2}$ is the bases of subspace $S_2, d_{s_{k-1}+1}, \dots, d_n$ is the bases of subspace S_k respectively. For the point in P can't be expressed as the linear combination of d_i , we apply a transformation on P and denote the result as V . We build EMST on P (denoted as T), then extract all adjacent points p, q in T , and store the vector $p-q$ in V . Most of the vectors in V can be represented as sparse linear combination of the columns of D . But there are some exceptions. The vectors generated from the non-manifold points or sharp feature points will not fit the local linear model due to the limitation of EMST. As the amount of these vectors is relatively small, we can just take these vectors as noise. Constraint in dictionary learning problem $x = D\alpha$ can be replaced by $x \approx D\alpha$ to handle noise or error in samples. With Lagrange multiplier method, the sparse local structure recognition model can be written as:

$$\min_{\alpha} \|\alpha\|_0 + \lambda \|V - D\alpha\|_2^2 \quad (2)$$

For a given n , we can get d_1, \dots, d_n through any dictionary learning algorithm, as an example of how this might work in practice, see Fig 2. In this paper, we use SPAMS to solve such problems.



Fig. 2. Point cloud sparse decomposition. For a simple 2D example, n has been set to 3. Learned D has been shown as d_1, d_2, d_3 . Direction vector using the points in ℓ_i decompose on D will only has a non-zero coefficient on d_i .

3.1 Extract Bases of Subspaces by Iterative Algorithm

In the previous section, we have seen that, for a suitable given n , bases of subspaces can be found through P(2). But in practice, it is hard to know how many columns D should have. In this section, we will describe the iterative algorithm which can automatically choose a proper n .

In 3D space, S_i may present as line or plane, both of which are discussed separately in the following sections.

Consider S_i is a line, α will be sparsest when the bases of S_i have only one entry which corresponds to the slope of S_i . If we set $n_i \gg 1$ for S_i , the learned bases will be zero entry or slightly different with the slope of S_i duo to noise. In this case, we can discard the zero entries and merge two columns if the angle between them is small enough, which means they represent the same structure.

If S_i is a plane, the learned bases will be different when the points' distribution on that plane changes. For example, if the points are sampled uniformly from a square, the learned bases will segment the unit ball equally. If sampling process is changed from uniform to stochastic, the bases changed. But whatever happens, the bases must lay on the same plane. From another side, the basis of S_i usually will be redundant, the number of basis is determined both by the parameter λ in P(2) and the learning samples in V. For a given point set and λ , if we iteratively apply dictionary learning algorithm on it, try to reduce the number of basis by merge the vectors which are approximate parallel, the algorithm will converge at some number n_i .

Based on the discussion above, we propose an iterative algorithm to extract the bases of subspaces. The basic idea of this algorithm is: Firstly, we use a large n to learn an overcomplete dictionary. Then, we analysis the columns of D, discard the empty columns and merge the similar ones, get $n_2 \leq n$ for next iteration. Thirdly, we recursively train the dictionary with new value of n until $n_{k+1} = n_k$. After these steps, we can acquire the desired bases of subspaces. An example of how this algorithm works can be found in Fig 3.

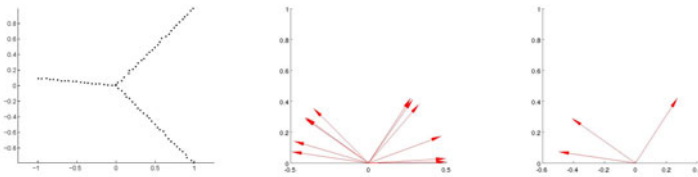


Fig. 3. An example of iterative method with 3 coplanar intersection lines. After 5 iterations, we get the exactly three bases for input point cloud.

3.2 Identify Subspace

In this section, we discuss how to identify the dimensions of the subspaces and the bases belong to them. Identify subspace have two main problems:

1. How do we identify plane? The bases for a plane will lay on the same plane, so we can check the coplanarity of d_i in R^2 manifold. If some bases are found coplanar, it's possible that they are overcomplete bases of a plane.
2. How do we distinguish the coplanar lines with plane? To separate the coplanar lines from plane, we analysis the sparse coefficient vector α . Samples decomposition on D can be written as $x_i = d_1\alpha_{i1} + \dots + d_n\alpha_{in}$.

Assume the set $D_{sub} : \{d_a, \dots, d_b\}$ are coplanar, we first find the samples v_{sub} which use these columns, then check the a th element to the b th one in α : if the diversity is one, then D_{sub} stands for lines, otherwise plane.

Based on the above discussion, once we get $\{d_1, \dots, d_n\}$, we can analysis their coplanarity in R^2 manifold and the coefficient on these columns to extract planes. The columns left in D represent lines.

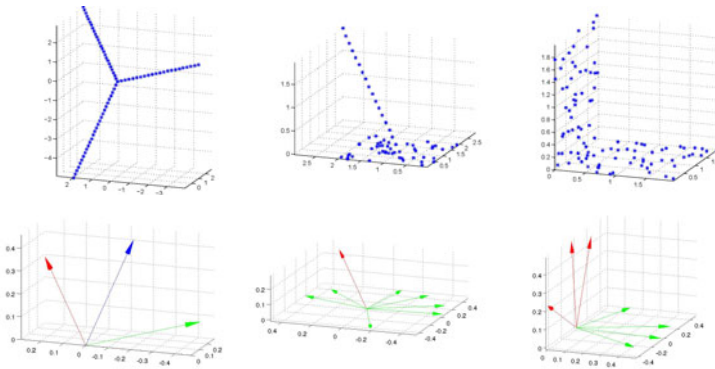


Fig. 4. Experiment results. The topmost row shows the original point set. The bottom row shows the obtained results. The first column shows 3 non-coplanar intersection lines; the middle column shows result of line intersection with plane; the last column shows the situation of 2 intersection planes.

4 Experiments and Analysis

We present results obtained with the proposed iterative methods. We show examples of 3 non-coplanar intersection lines, line intersection with plane, and 2 intersection planes. Experiments are running in matlab. Results are shown in Table 1 and Fig 4.

Table 1. Experiment results

	λ	Iteration number	Obj function	Total running time(sec)
Exp1	0.05	4	0.018262	2.24
Exp2	0.15	7	0.018862	2.85
Exp3	0.4	7	0.019427	3.01

We can see that, our method can converge rapidly after a few iterations, and the errors between learned bases and the real ones are small. Solutions for P(3) often has higher time complexity, but due to the small point size in local region, the total running time is acceptable.

5 Conclusions

In our work, we model the local structure of point cloud as the union of subspaces. An iterative algorithm has been proposed, it can learn the bases of subspaces which are

approximated as linear. Once get the bases of subspaces, we use a post analysis on the coplanarity of these vectors and the projected coefficient to identify the subspaces. Examples proved the efficiency of this method in both 2D and 3D cases. With the local structure of point cloud, one can use it to do robust normal estimation, dimension detection et al.

References

1. Huang, H., Li, D., Zhang, H., Ascher, U., Cohen-Or, D.: Consolidation of unorganized point clouds for surface reconstruction. In: *ACM Transactions on Graphics (TOG)*, vol. 28, p. 176. ACM (2009)
2. Avron, H., Sharf, A., Greif, C.: And D. Cohen-Or. 11-sparse Reconstruction of Sharp Point Set Surfaces. Submitted to *ACM Transactions on Graphics* (2009)
3. Fleishman, S., Drori, I., Cohen-Or, D.: Bilateral mesh denoising. *ACM Transactions on Graphics (TOG)* 22(3), 950–953 (2003)
4. Lewicki, M.S., Sejnowski, T.J.: Learning overcomplete representations. *Neural computation* 12(2), 337–365 (2000)
5. Vidal, R., Ma, Y., Sastry, S.: Generalized principal component analysis (gpca). In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1945–1959 (2005)
6. Mairal, J., Bach, F., Ponce, J., Sapiro, G.: Online dictionary learning for sparse coding. In: *Proceedings of the 26th Annual International Conference on Machine Learning*, pp. 689–696. ACM (2009)
7. Aharon, M., Elad, M.: Sparse and redundant modeling of image content using an image-signature-dictionary. *SIAM Journal on Imaging Sciences* 1(3), 228–247 (2008)
8. Mairal, J., Bach, F., Ponce, J.: Task-driven dictionary learning. *Arxiv preprint arXiv 1009.5358* (2010)
9. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: Design of dictionaries for sparse representation. *Proceedings of SPARS* 5 (2005)
10. Engan, K., Aase, S.O., Husøy, J.H.: Multi-frame compression: Theory and design. *Signal Processing* 80(10), 2121–2140 (2000)

Pyroelectric Infrared Sensors for Human Identification Using Non-negative Matrix Factorization and BP Neural Network

Ning Zhao, Fangmin Li, and Sheng Wang

School of Information Engineering, Wuhan University of Technology, Wuhan, 430070, China
znwhlg@126.com, zhaoning@whut.edu.cn

Abstract. Pyroelectric Infrared (PIR) sensors are excellent devices for wireless sensor network due to its characteristics of low-cost and low-power. PIR sensors are widely used to establish simple but reliable system for detecting targets or triggering alarms. However, processing numerous output data from PIR sensors and correcting the high false generated in the process of classifying and identifying of human targets limit the application scope of PIR sensors. In this paper, a feature extraction and sensor data fusion method to detect and recognize multiple human targets moving in a detection area are presented. Simulation results shows that such approach can reduce computational requirement which indicates that PIR sensors can be used as wireless sensor nodes with limited resources. Additionally, when using the BP neural network, the system can achieve 96% correct identification of individual target data and 90% correct classification of multiple targets mixed data as well.

Keywords: Pyroelectric Infrared (PIR) sensors, feature extraction, identification.

1 Introduction

Pyroelectric Infrared (PIR) sensors are small, low-power, low-cost, passive sensors used to detect human bodies which are not at thermal equilibrium with the environment. They can capture incident radiation by means of the change in their temperature. These sensors are typically used in commercial applications to detect the presence of individuals and/or trigger alarms. PIR sensors are also used in much more complex applications such as thermal imaging [1], radiometry [2], thermometers [3], and biometry [4], etc.

In this paper, we present methods to identify multiple individuals using wireless PIR sensors. These methods need low computational power, which are suitable for a parallel implementation of wireless nodes that have limited resources, and further reduce the obtrusiveness and cost since no wires are needed.

The rest of the paper is organized as follows. In section 2, we introduce PIR sensors and their working principles. Section 3 describes the feature extraction of human targets. The data fusion and the BP neural network training are presented in section 4 and section 5. Finally, we present experimental results and conclude the paper.

2 PIR Sensors and Simulation Signals

Pyroelectricity is the electrical response of a polar, dielectric material to a change in its temperature. A pyroelectric element converts incident IR flux into an electrical output by two steps: the absorbing layer transforms the radiation flux change into a change in temperature; and the pyroelectric element performs a thermal to electrical conversion [5].

The PIR sensors are used in conjunction with Fresnel lenses to augment and shape their FoV [6]. Typically, an array of Fresnel lenses is used to divide the PIR sensor FoV into several optically separated cones. Generally, PIR sensor has two detection elements with opposite polarity, then one Fresnel lens can form two light areas and one dark area, i.e., $2n$ light areas and $2n - 1$ dark areas for n lenses; thereby, the view of the detection range of a detector is divided into $4n - 1$ light and dark areas.

The process that a human target walks through a complete light or shade detecting area can be divided into 4 stages with 4 subfigures. From the first stage to the forth one, the human target walks through alternate areas among the light detection zones and the darkness ones successively; and the output signal of PIR sensors change positively and negatively along with some critical parameters such as target’s speed, the height and the effective radiation proportion of human targets. If we set speed, width, height and proportion as v, w, h and y , the relationship of these three parameters is given by

$$\left\{ \begin{array}{l} y = h \times v \times t, 0 < t \leq \frac{w}{v} \\ y = h \times w, \frac{w}{v} < t \leq \frac{a}{v} \\ y = h \times (w - v \times t), \frac{a}{v} < t \leq \frac{a + w}{v} \\ y = 0, \frac{a + w}{v} < t \leq \frac{2a + w}{v} \end{array} \right. \tag{1}$$

$$\left\{ \begin{array}{l} y = -h \times v \times t, \frac{w + 2a}{v} < t \leq \frac{2w + 2a}{v} \\ y = -h \times w, \frac{2w + 2a}{v} < t \leq \frac{3a + w}{v} \\ y = -h \times (w - v \times t), \frac{3a + w}{v} < t \leq \frac{3a + 2w}{v} \\ y = 0, \frac{3a + 2w}{v} < t \leq \frac{4a + 2w}{v} \end{array} \right. \tag{2}$$

In this paper, we assume that there are 5 persons walking through the exploration area at the same time. The human target can be modeled by a vertical cuboids radiation source approximately. According to Chinese human body size standards: the average height of Chinese men is 1.69m, the average shoulder width is 0.42m and the average width of body side is 0.2m [7]. For the convenience of experiments, we hypothesize the height of 5 persons is 1.5m, 1.6m, 1.7m, 1.8m, 1.9m respectively and the width of all 5 persons take mean value as 0.62m. So the simulation signals that 5 individuals walk through the sensor area with different speeds are shown in Fig.1

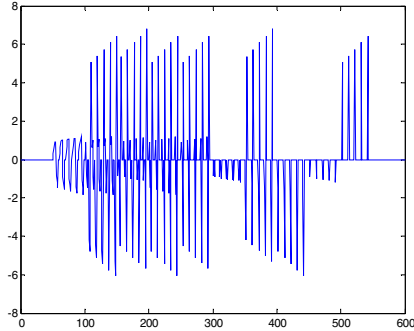
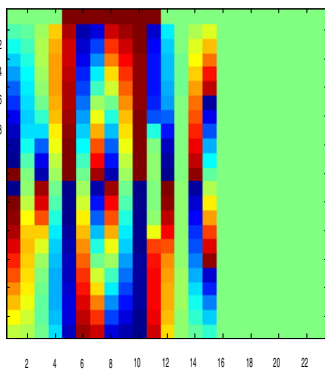


Fig. 1. Time domain simulation signals of five human bodies

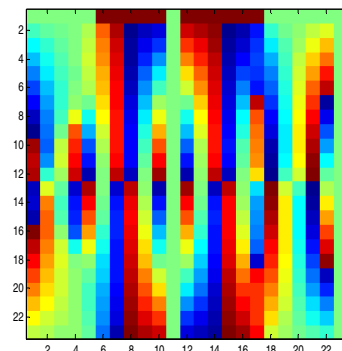
3 PIR Frequency Domain Feature Extractions

PIR signal is a kind of random signal. For a random signal, it could not use Fourier transform to analyse, yet it could not be described by a given accurate mathematical expression. In general, the autocorrelation functions, which are Fourier transform pairs with power spectral density, can express specific statistical average volume of random signal completely; so, it can be used to characterize the power spectral density statistical average spectrum characteristics.

Fourier transform and spectral estimation have high capability of characterizing local features in the frequency domain. In order to effectively solve the problem of localized analysis to non-stationary signal, we adopt Short Time Fourier Transform (STFT) algorithm proposed by Gabor [8]. It cuts linear time signal $x(\tau)$ into pieces with special window function $h(\tau)$ which is a continuous function with compact



Height $h=1.7\text{m}$, speed $v=1.2\text{m/s}$



Height $h=1.5\text{m}$, speed $v=0.8\text{m/s}$

Fig. 2. This figure is two feature spectrograms of two human targets with different heights moving at different speeds

support, as we define $h(\tau)$ as Hamming window, we can get the human target features spectrum of different persons walking in different speeds. In Fig.2, We can clearly see that the distinction of these two spectrograms is obvious.

4 High-Dimensional Data Fusion

After transforming by STFT, multiple human target signals detected by wireless PIR network nodes are high-dimensional aliasing frequency domain signals. The time and energy consumption will increase if those data are operated directly by PIR nodes without any processing, which is not suitable for battery-powered wireless nodes. Therefore, we intend to use non-negative matrix factorization (NMF) algorithm to fuse aliasing signals data before classification and identification.

NMF algorithm is the research finding about non-negative matrix proposed by Lee and Seung [9], which was published in *Nature Journal*, 1999. In the article, they proposed an idea of matrix decomposition under some constraint conditions where the decomposed matrix can not appear negative value and can be described as: known as a $n \times m$ matrix V , each column of which contains n non-negative pixel values of one of the m human body features. Then, the method constructs approximate factorizations of the form $V \approx W \times H$, or described as

$$V_{i\mu} \approx (WH)_{i\mu} = \sum_{a=1}^r W_{ia} H_{a\mu} \quad (3)$$

The r columns of W are called basis images. Each column of H is called an encoding and is in one-to-one correspondence with a human target in V . An encoding consists of the coefficients by which a face is represented with a linear combination of basic data. The dimensions of the matrix factors W and H are $n \times r$ and $r \times m$, respectively. The rank r of the factorization is generally chosen so that $(n+m)r < nm$, and the product WH can be regarded as a compressed form of the data in V . The dimension of W and H will be smaller than the original matrix V , which can reduce the data dimension and then maximize data storage capacity.

We put mixed feature frequency domain signals that have been transformed by STFT according to the original data matrix V which is 600×600 . By using NMF iterative algorithm we can get both the sparse basis matrix W , which is 600×20 , and encoding matrix H . Compared with the original data matrix V , W not only is low dimension that will be pretty applicable for PIR wireless nodes, but also has not lost the feature of the five human targets so as to classify and recognize individual human target from them expediently.

5 Classification and Identification of Human Targets

5.1 Design of BP Neural Network

We plan to adopt three-layer BP neural network in order to distinguish single individual from the mixed five human targets' STFT frequency domain signals. In this paper,

we use STFT basis matrix W as the BP network input vector. Because of the W , we select the number of input neurons of such BP network input layer is 20.

The dimensions of output layer which represents the number of neurons of output layer are generally determined according to the user's requirements. If the BP neural network is used as a classification with m different categories, then the output layer neurons are usually m or $\log_2 m$. Therefore, for classifying five human targets, the number of neurons of such BP network output layer is set to 3. We use binary code 000, 001, 010, 100, 111 representing the first, second, third, fourth and fifth human target, respectively. Meanwhile we select the number of hidden layer neurons is 41.

After STFT and NMF iterative operation, we can get each 120×20 feature frequency spectrum matrix severally about the five persons. Then, we take five 100×20 sub-matrixes from five human targets feature matrixes respectively to form a new 500×20 matrix as training samples of the BP neural network. And, the rest data of each feature matrix are test samples in order to ensure the training samples and test samples are irrelevant.

In addition, we take $1 \times e^{-3}$ as BP network error threshold that network training will be terminated when the training error has reached such threshold value. The *tansig* function and *trainidx* function are chosen as transfer function and training function of the BP network respectively and the initial value of the network choose the default value. After being trained for several times with those above parameters, the network will be continuously modified in terms of some network weights and thresholds until the performance of the network achieve stable state.

5.2 Classification and Identification of the BP Network Test Experience and Results

Experience 1: We make the testing samples which are selected in section 5.1 as input vectors of the BP neural network just trained, we can clearly get from the result data that the classification and identification rate of human target using this BP network is about 96%.

Experience 2: We make a 100×20 matrix which is randomly selected from that 600×20 overlapped feature frequency domain data matrix of five human targets as input vectors of the BP neural network just trained, The results show that such BP neural network has the capability to classify and distinguish five human targets from overlapped frequency domain data matrix including all features of the five targets. Besides, the correct rate is nearly 90%.

6 Conclusion

In this paper, we use STFT to extract frequency domain features of multiple human targets from time domain signals and NMF iterative algorithm to fuse the high-dimensional original signal data. With these strategies, the computation and power consumption of sensor nodes can be reduced, making such nodes are more suitable for wireless PIR sensor network; besides, the trained BP neural network is relatively

simple and easy to achieve. The experiences' results prove that the classification and identification data can meet application requirements.

Acknowledgments. This work is supported by the National Science Foundation of China, under Grant No.61170090.

References

1. Astheimer, R.W., Schwarz, F.: Thermal imaging using pyroelectric detectors. *Appl. Opt.* 7, 1687–1696 (1968)
2. Pradhan, M.M., Gard, P.K.: Pyroelectric null detector for absolute radiometry. *Appl. Opt.* 21(24), 4456–4458 (1982)
3. Tsai, C.F., Young, M.S.: Pyroelectric infrared sensors-based thermometer for monitoring indoor objects. *Rev. Sci. Instrum.* 74, 5267–5273 (2003)
4. Fang, J.S., Hao, Q., Brady, D.J., Guenther, D.D., Hsu, K.Y.: Real-time human identification using a pyroelectric infrared detector array and hidden Markov models. *Optics Express* 14(15), 6643–6658 (2006)
5. Milde, G., Hausler, C., Gerlach, G., Bahr, H.-A., Balke, H.: 3-D modeling of pyroelectric sensor arrays Part II: Modulation transfer function. *J. of IEEE Sensors* 8(12), 2088–2094 (2008)
6. Cirino, G.A., Barcellos, R., Berezki, A., Morato, S.P., Neto, L.G.: Design, fabrication and characterization of Fresnel lens array with spatial filtering for passive infrared motion sensors. In: *Proc. Photon. North Int. Conf. Appl. Photon. Technol.*, pp. 1–12 (2006)
7. Kakuta, N., Yokoyama, S., Nakamura, M.: Estimation of radiation heat transfer using a geometric human model. *IEEE. Trans. Biomed. Eng.* 48, 324–331 (2001)
8. Gabor, D.: *Theory of Communication*. *J. Inst. Elect. Eng. London* 93, 429–457 (1946) (assigned 8/25/04)
9. Lee, D.D., Seung, H.S.: Learning the parts of objects by non-negative matrix factorization. *Nature* 401(6755), 788–791 (1999)

Bimodal Emotion Recognition Based on Speech Signals and Facial Expression

Binbin Tu and Fengqin Yu

School of Internet of Things Engineering, Jiangnan University,
Wuxi 214122 China
tbbice@126.com

Abstract. Voice signals and facial expression changes are synchronized under the different emotions, the recognition algorithm based audio-visual feature fusion is proposed to identify emotional states more accurately. Prosodic features were extracted for speech emotional features, and local Gabor binary patterns were adopted for facial expression features. Two types of features were modeled with SVM respectively to obtain the probabilities of anger, disgust, fear, happiness, sadness and surprise, and then fused the probabilities to gain the final decision. Simulation results demonstrate that the average recognition rates of the single modal classifier based on speech signals and based on facial expression reach 60% and 57% respectively, while the multimodal classifier with the feature fusion of speech signals and facial expression achieves 72%.

Keywords: speech emotion recognition, facial expression, local Gabor binary patterns, support vector machine, fusion.

1 Introduction

The voice contains the rich information of emotion, while the face as the most important external feature can transmit much non-verbal information to enhance, understand or express emotion [1]. And psychologist Mehrabian gave a emotional formula as feelings show = 7%utterance+ 38%voice+ 55%countenance. In recent years, the emotion recognition has developed towards multi-modal. Professor Picard of MIT Media Lab extracted the multi-modal features based on facial expressions, head gestures and so on to monitor real-time emotional states of the students in the learning [2].

The emotional information of voice is contained in the changes of the acoustic parameters including prosodic parameters and spectral parameters [3]. LPCC reflects the characteristics of channel physiological structure, while MFCC reflects the nonlinearity of people's auditory frequency [4]. Local Binary Pattern (LBP) proposed by T. Ojala is an effective texture description operator and can measure and extract the texture information in the local neighborhood of gray-scale image [5]. T. Ahone firstly introduced the theory of LBP to describe the face image areas face, divided face images into several regions, extracted the LBP texture features from them and achieved good results [6]. And Gabor wavelet transform can describe the multi-scale, multi-direction local features of the gray-scale image, imitate the contours of single-cell receptive in the cerebral cortex and capture the prominent visual properties [7].

Therefore, Local Gabor Binary Patterns (LGBP) can extract multi-directional, multi-scale local image features using Gabor transform, and then encode these features with LBP operator to effectively distinguish the different facial expression images.

The remainder of this paper is organized as follows. Section 2 describes the algorithm theory of LGBP and the fusion rules of classifiers. Section 3 presents simulation steps of extracting features from bimodal emotion signals and recognizing. Section 4 obtains the results based on speech signals and facial expression. Conclusions and discussions are given in section 5.

2 Algorithm Theory

2.1 LBP

LBP is originally used for texture analysis problems, which is successfully applied to the face recognition area in recent years. The principle of LBP is to calculate the binary sequence in the gray varying information of each pixel in the image, then encode the relation of the binary sequence to form LBP. So LBP features have much gray variation information. A pixel in an image denoted as $g(x_c, y_c)$, the original LBP operator labels the pixels of an image by threshold the 3×3 neighborhood of each pixel g_0, g_1, \dots, g_7 with the center value g_c , defining the texture of the local area as $T = t(g_c, g_0, g_1, \dots, g_7)$, then binary processing as follows:

$$T \approx t(s(g_0 - g_c), s(g_1 - g_c), \dots, s(g_7 - g_c)) \tag{1}$$

Where $s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$. Describing the spatial structure of local texture characteristics of LBP expressed as

$$LBP_{8,1}(x_c, y_c) = \sum_{i=0}^7 s(g_i - g_c) 2^i \tag{2}$$

Where $LBP_{8,1}$ is LBP operator based on a circularly symmetric neighbor set of 8 members on a circle of radius 1, the LBP operator in Figure 1, a happy face image corresponding to the coded image shown in Figure 2.

With the increasing of the sampling points, the type of binary mode to increase dramatically, while the number of texture features dimension is larger, it is not conducive to classify. So T.Ojala proposed "uniform" patterns to improve the original LBP model, when the LBP corresponding to the binary number which changes from 0 to 1 or from 1 to 0, and its changes are no more than two times, U value is at most 2,

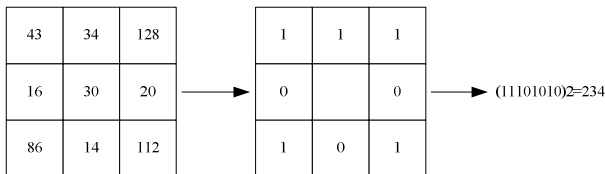


Fig. 1. LBP operator



Fig. 2. A happy face image corresponding to the LBP image

denoted as $P_{8,1}^{\mu 2}$, for example, patterns 00000000, 11111111, 11000011 are “uniform” patterns, defined as:

$$P_{8,1}^{\mu 2} = \begin{cases} \sum_{i=0}^7 s(g_i - g_c)2^i, U \leq 2 \\ 8 + 1, & \textit{otherwise} \end{cases} \tag{3}$$

When using the (8, 1) field, the “uniform” patterns account for 90% of all patterns, which can effectively describe most of the image texture features and reduce the number of features greatly. Therefore, this paper takes (8, 1) field and “uniform” patterns of facial expression features.

2.2 LGBP Operator

LGBP operator is to combine the amplitude of the Gabor wavelet features and LBP encoding. Gabor features are widely used in visual understanding information, the transform coefficients have well visual characteristics and are sensitive to the edge of the images. Gabor filters facilitate to adjust the direction, the base frequency bandwidth, and have a higher resolving power in time and frequency domain, two-dimensional Gabor wavelet kernel function is given:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|}{\sigma^2} e^{\left(-\|k_{\mu,v}\|^2 \|z\|^2 / 2\sigma^2\right)} \left[e^{ik_{\mu,v}z} - e^{-\sigma^2/2} \right] \tag{4}$$

Where $z=(x, y)$ is image pixel, norm operation denotes $\|\bullet\|$, μ and v represent the direction and scale of Gabor filters respectively. $k_{\mu,v} = k_v e^{i\varphi_\mu}$, $k_v = 2^{-(v+2)/2} \pi$, $\varphi_\mu = \pi\mu / k$, k represent total number of wavelets directions, k_v is kernel frequency, and σ is wavelet filter bandwidth.

Gabor features of face images are gained by the convolution of facial images and Gabor filters, $f(x, y)$ is the gray distribution of facial images, and Gabor features as:

$$G(x, y, \mu, v) = f(x, y) * \psi_{\mu,v}(z) \tag{5}$$

Eight-direction and five-scale Gabor filters are adopted in this paper, $\mu = 0, 1, \dots, 7, v = 0, 1, 2, 3, 4$. LGBP is gained by the formula (2) and (5):

$$LGBP = \sum_{i=0}^7 s(G_i(x, y, \mu, v) - G_c(x, y, \mu, v))2^i \tag{6}$$

2.3 Fusion Rule

Kittler proposed the theoretical framework based on minimum error rate Bayesian classifier[8].Considering a pattern recognition problem where pattern Z is to be assigned to one of the C possible emotion classes $\Omega = \{\omega_1, \omega_2 \dots \omega_c\}$, extracted R group features $f_1, f_2 \dots f_R$ from a sample, we have R classifiers, and each classifier corresponding to the feature f_R , R group features are modeled by the SVM, then gained posterior probabilities $P(\omega_i | f_i) \sim P(\omega_i | f_R)$ of the i -th emotion. Take the maximum of posterior probabilities of the corresponding categories as the fusion result:

$$r(Z) = \arg \max F [P(\omega_i | f_1), \dots, P(\omega_i | f_R)] \tag{7}$$

Function F represent fusion rule, including sum rule (8) and product rule (9).

$$F [P(\omega_i | f_1), \dots, P(\omega_i | f_R)] = \sum_{j=1}^R P(\omega_i | f_j) \tag{8}$$

$$F [P(\omega_i | f_1), \dots, P(\omega_i | f_R)] = \prod_{j=1}^R P(\omega_i | f_j) \tag{9}$$

3 Simulation Steps

Emotion recognition based on speech and facial expression was carried out as follows:

- (1) Pre-process the emotional speech samples, including pre-emphasis and window them into frames. Frame size was set to 256 points, and frame shift 128, while the hamming win was chosen;
- (2) For each frame signal, compute its FFT to gain the energy spectrum which will then input into Mel filter group to extract 12-dimensional MFCC;
- (3) Compute the average fundamental frequency, short-time average energy, and 12-dimensional LPCC of each frame signal;
- (4) Read the facial expression images, and pre-process them with geometry normalize and grayscale normalize, each image was normalized into 123×123, and maintain the eyes, eyebrows, mouth of the face image in the same location;
- (5) Make eight-direction, five-scale Gabor wavelet decomposition for the images above, and take $\mu = 0, 1, \dots, 7$, $\nu = 0, 1, 2, 3, 4$, $\varphi_\mu = 0, \frac{1}{8}\pi ; \dots, \frac{7}{8}\pi$, $\sigma = 2\pi$. Then achieve LBP code to calculate the LGBP features of each sub-images;
- (6) Input speech features which include fundamental frequency, energy, LPCC, MFCC and image features which contain the LGBP into SVM respectively, finally the recognition rates of six emotions were obtained;
- (7) Get all emotions' posterior probability based on speech features and image features respectively;
- (8) Fuse the posterior probability of two features according to sum and product rules, calculate the recognition rates corresponding to the six emotions again.

4 Simulation and Results

The eINTERFACE'05 audio-visual emotion database was adopted in this paper, including six emotions of happiness, sadness, surprise, anger, disgust and fear, containing 42 subjects, recorded by the people of 14 different nationalities. The speech and facial samples were cut from the video files by 16 kHz sampling rate, 16 bit sampling precision and 25 frames/sec image sampling rate, we selected 900 samples of voice and facial samples of five persons as the experimental test sample.

Extracted pitch frequency, short-time average energy, 12-dimensional LPCC and MFCC from speech signals, then using SVM to identify six emotions, the results is shown in Table 1, average recognition rate is 60% based on speech signals, disgust and sadness are the most confusing emotion, surprise recognition rate is highest.

LGBP features are extracted from facial expression images, six kinds of emotional recognition results are shown in Table 2, the average recognition rate is 57%, when only using facial expression to identify emotions, fear and surprise are the most confusing emotions, happiness achieve the highest recognition rate.

Table 1. The recognition rate of the six emotions based on speech signal (%)

emotion	anger	disgust	fear	happiness	sadness	surprise
anger	60	17	0	0	23	0
disgust	0	53	7	0	33	7
fear	27	0	40	13	0	20
happiness	20	0	0	60	0	20
sadness	0	13	20	0	67	0
surprise	0	0	13	7	0	80

Table 2. The recognition rate of the six emotions based on facial images (%)

emotion	anger	disgust	fear	happiness	sadness	surprise
anger	60	13	0	27	0	0
disgust	0	40	6	0	27	27
fear	0	0	53	0	13	33
happiness	20	0	0	80	0	0
sadness	7	13	0	13	67	0
surprise	27	0	20	0	13	40

Using sum and product rules to fuse the features of speech signal and facial expression, the results of sum rule are better in table 3, the average recognition rate reach 72%, which is improved significantly. Anger and sadness recognition rate increased to 80%, fear and surprise confusing error rate have been reduced.

Table 3. The recognition rate based on fusion of speech signal and facial images (%)

emotion	anger	disgust	fear	happiness	sadness	surprise
anger	80	0	0	13	7	0
disgust	0	53	20	0	27	0
fear	0	0	67	0	20	13
happiness	13	0	0	87	0	0
sadness	0	13	0	7	80	0
surprise	7	0	13	0	13	67

5 Conclusion

While most emotion recognition researches rely on single-mode data, speech and image features were both extracted for bimodal emotion recognition in this paper. With fusion of speech and facial image features, an effective emotion recognition approach based on two kind biological signals was implemented. As is shown in the experiments, recognition rate after feature fusing was improved significantly.

Acknowledgments. This paper was supported by the National Natural Science Foundation of China No.61075008. The authors would like to thank Martin who provided eNTERFACE'05 audio-visual emotion database.

References

1. Jinjing, X., Yiqiang, C., Junfa, L.: Multi-expression Facial Animation based on Speech Emotion Recognition. *Journal of Computer-aided Design & Computer Graphics* 20(4), 520–525 (2008)
2. Kapoor, A., Picard, R.W.: Multimodal Affect Recognition in Learning Environments. In: *Proc. of the 13th Annual International Conference on Multimedia*, Singapore, pp. 677–682 (2005)
3. Danning, J., Lianhong, C.: Speech Emotion Recognition using Acoustic Features. *J. Tsinghua Univ (Sci. & Tech.)* 46(1), 86–89 (2006)
4. Koolagudi, S.G., Nandy, S., Rao, K.S.: Spectral Features for Emotion Classification. In: *2009 IEEE International Advance Computing Conference*, Patiala, pp. 1292–1296 (2009)
5. Ojala, T., Pietikainen, M., Maenpaa, T.: Multi-resolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7), 971–987 (2002)
6. Ahonen, T., Hadid, A., Pietikainen, M.: Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(12), 2037–2041 (2006)
7. Wenchao, Z., Shiguang, S., Hongming, Z.: Histogram Sequence of Local Gabor Binary Pattern for Face Description and Identification. *Journal of Software* 17(12), 2508–2517 (2006)
8. Kittler, J., Hatef, M., Duin, R.P.: On Combining Classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(3), 226–239 (1998)

Memory-Based Multi-camera Handover with Non-overlapping Fields of View

Xiaoyan Sun^{1,2}, Faliang Chang¹, and Jiangbao Li¹

¹ School of Control Science and Engineering, Shandong University

² School of Computer Science and Technology, Shandong Jianzhu University
Jinan, Shandong Province, 250061, China
sxy@sdjzu.edu.cn

Abstract. Object tracking is an important task within the field of computer vision, and the multi-camera tracking with disjoint view is more applicable. This paper focus on introducing human memory mechanism into multi-camera human tracking problem that the Field Of View (FOV) of cameras are not necessarily overlapping, and proposing a camera handover scheme based on memory. In the modeling process, every target goes through transmission and storage of three spaces: sensory memory, short-term memory and long-term memory. After learning, memory-based target handoff can remember target appeared earlier, and when faced with similar goals, it can extract and activate the target in memory in time, so that it can quickly achieve object handoff between different camera tracking. Preliminary experiments show that this scheme is effect in camera handover and multi-camera human tracking.

Keywords: Object tracking, Camera handover, Human tracking, Human memory.

1 Introduction

Object tracking is an important task within the field of computer vision, and it can be used in motion-based recognition, automated surveillance, human-computer interaction, traffic monitoring, vehicle navigation, etc. With the limited Field Of View (FOV) of single camera, it is necessary to use multiple, distributed cameras to monitor a wide area completely and VSAM [2] and ADVISOR [3] are examples. Several papers have previously discussed multi-camera tracking with overlapping FOVs, such as [5] and [6]. However, in practice, it is hard to expect that the FOVs of cameras are always overlapping to each other, so tracking objects across non-overlapping cameras is more applicable. Jinman Kang [8] defines a spatio-temporal Joint Probability Data Association Filter (JPDAF) for integrating multiple cues including appearance and motion. Kettner and Zabih [9] employed a Bayesian formulation of the problem to reconstruct the paths of objects across multiple cameras, and transformed it into a linear-programming problem to establish correspondence. O. Javed etc. [10] combines multiple cues including inter-camera time intervals, location of exit/entrances, and velocities of objects to constrain correspondences in a Bayesian framework for tracking.

In this paper, we focus on introducing human memory mechanism into multi-camera human tracking problem that the FOV of cameras are not necessarily overlapping, and proposing a camera handover scheme based on memory. In the modeling process, every target goes through transmission and storage of three spaces: sensory memory, short-term memory and long-term memory. After learning, memory-based target handoff can remember target appeared earlier, and when faced with similar goals, it can extract and activate the target in memory in time, so that it can quickly achieve the object handoff between different camera tracking.

The reminder of the paper is organized as follows: Section 2 introduces the human memory mechanism. We elaborate our memory-based multi-camera handover with disjoint view in section 3. Section 4 illustrates our experimental results and Section 5 concludes this paper.

2 Human Memory Model

As is well known, no matter how the scene changes, man can remember the perceived scene and objects in a certain period of time, and can easily distinguish the target and background. According to cognitive psychology, why human is able to do this is closely related to the human race has a functional and powerful memory system. The things that people have seen and experienced will go through the treatment of memory system. When perceiving new things, the memory information associated with the things will be extracted, thereby speeding up the recognition process. Modern cognitive psychology considers the human memory model as in Figure 1.

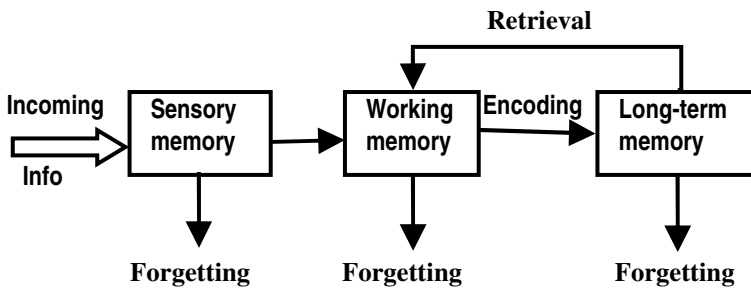


Fig. 1. Human memory model

As shown in Fig. 1, there are generally three types of memory: sensory memory, short-term memory (working memory) and long-term memory, and there are three main activities related to long term memory: storage, deletion and retrieval. The information not frequently used or extracted will be lost from memory, which is forgetting.

1). Sensory Memory

The sensory memories act as buffers for stimuli received through the senses. Information is passed from sensory memory into short-term memory by attention, thereby filtering the stimuli to only those which are of interest at a given time.

2). Short-Term Memory (Working Memory)

Short-term memory is used to make decisions. Information stored in short-term memory including the new information obtained by sensory memory, the information stored in short-term memory which have been processed, and the information extracted from long-term memory. Short-term memory can be seen as a complex system which store and process information.

3). Long-Term Memory

Long-term memory is a large and complex information database. Information in working memory is transferred to it after a few seconds. Information stored in long-term memory enables subject recall a variety of things have occurred, acknowledge model and thus solve problems.

3 Memory-Based Multi-camera Handover with Disjoint View

Based on the human memory model, we construct the multi-camera handover and human tracking system and the skeleton is show in Fig. 2.

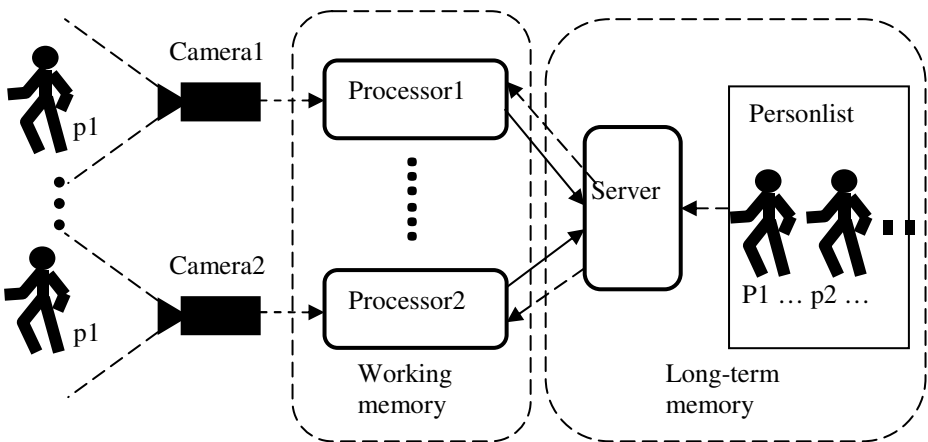


Fig. 2. The proposed memory-based multi-camera tracking system

In Fig. 2, two cameras are taking for example, and it can be extended to multiple cameras. Cameras represent the sensory memory, which take video sequences, and transport to working memory (that is processors); processors detect human targets and track them in video sequences from video cameras, compare with the models extracted from personlist on server to identify them, so they are in the place of working memory. Each processor can be a separate computer, or a thread in a process. Server is the long-term memory which stores the personlist, which recorded all target models have appeared, and it can be extended by manual input such as blue torso, black legs. Corresponding to the processor, the server can be independent computer, or it can be a same computer with the processors.

Once a person enters the view of a camera, the processor linked with this camera will check if the person is registered in the personlist got from server. If he or she is the one on the list, the processor keeps on tracking the person. If not, it determines that he or she is a new person entering the network of cameras. If the person last for an enough time, the processor will register a new human model on the personlist. Each camera can do like this. With these continued checks and tracking, the cameras that are distributed over a wide area can be used to make a continuous tracking of a person.

3.1 Foreground Detection

In order to tracking objects accurately, detecting foreground exactly is the first step of tracking system. The proposed system use adaptive Single Gauss Model [11] to conduct background subtraction for moving person detection. Whenever a new frame arrives, a change in pixel intensity is computed using (1) to classify background or foreground (moving persons).

$$|I_t - u_B^t| \leq n * \sigma_B^t \tag{1}$$

Here, u_B^t and σ_B^t are the mean and standard deviations of the background at time t , respectively, and I_t is the new pixel value.

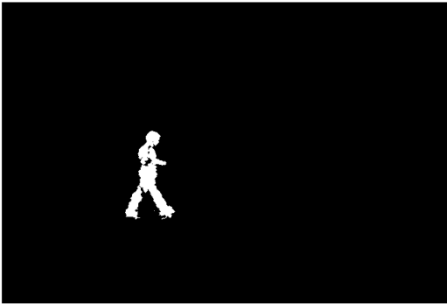


Fig. 3. Foreground detection



Fig. 4. Three Human blobs

If a pixel is classified to background, the adaptive background model is updated according to formula (2).

$$B_{t+1}(x, y) = \begin{cases} I_t(x, y) + \alpha[B_t(x, y) - I_t(x, y)], & \text{if } I_t(x, y) \in B_t(x, y) \\ I_t(x, y) + \beta[B_t(x, y) - I_t(x, y)], & \text{else.} \end{cases} \tag{2}$$

where, $0 < \alpha < 1, \beta \approx 1$

In (2), t is the time stamp, $I_t(x, y)$ and $B_t(x, y)$ are the pixel value and background model of t frame, $B_{t+1}(x, y)$ is the background model of $t+1$ frame. α and β are adapting rate of background pixel and non-background. With large α , this adaptive method can do with sudden illumination change well for it deals with background pixel and non-background pixel differently. In our experiment, we use $\alpha=0.5, \beta=0.99$.

In order to detection human blobs accurately, we use a connected-component algorithm to merge small blobs into neighboring blobs that share similar colors to

overcome over-segmentation generated by initial grouping. The filtering processing to remove blobs too small and blobs which obvious not meet the body proportion is following, and shadow remove based on HSV color space [12] also conducted. The human blobs will be got as show in Fig. 3.

3.2 Blob-Based Human Model

After foreground detection, each blob according to human proportion is divided into three parts and allocated to three human body parts (head, torso, legs). Generally, the height and width of human body has a certain ratio [13]. set the human body height H , then the shoulder height is about $0.81H$, lower limb length $0.52H$, and head high about $0.11H$. The separated three blobs are shown in Fig. 4.

Color is a commonly used feature in moving object identification and tracking. Although color is easily affected by illumination, it is more stable than other features when monitoring a wide area using multi-camera. So we use color information of blobs to construct the human model. Because HSV color space is not easily fouled by illumination than RGB color space, we target block from RGB space conversion to HSV space based on

$$\begin{aligned}
 V &= \max(R, G, B) \\
 S &= \begin{cases} (V - \min(R, G, B)) * \frac{255}{V} & \text{if } (V \neq 0) \\ 0 & \text{otherwise} \end{cases} \\
 H &= \begin{cases} (G - B) * 60/S & \text{if } (V = R) \\ 180 - (B - R) * \frac{60}{S} & \text{if } (V = G) \\ 240 + (R - G) * \frac{60}{S} & \text{if } (V = B) \end{cases}
 \end{aligned} \tag{3}$$

Then establishment the H component histogram, extract the highest frequent value of H as the main color of blobs and calculate its variance at the same time. So, the human model in personlist is:

$$P_i = \{H_i^{b1}, \sigma_i^{b1}, H_i^{b2}, \sigma_i^{b2}, H_i^{b3}, \sigma_i^{b3}\}, i = 1, 2, \dots, n \tag{4}$$

3.3 Camera Handoff

When a person enters FOV of a camera, human body detection will conducted first. Once the foreground blob is determined as human, the blob will be divided into three parts, and the blob-based model will be calculated. At the same time, working memory extract the personlist from long-term memory, use (5) to calculate the distance of human blobs with personlist models.

$$D_i = \sum_{j=1}^3 |H - H_j^{bj}| / \sigma_j^{bj} \tag{5}$$

Then choose the minimum of D_i as the best fitness person.

If all D_i are fitted (6), the person entered is judged as a new person. If this new person is continuously detected for N frames, add its model to server's personlist, and forget the longest not- used person model.

$$D_i > T_D \tag{6}$$

Where, T_D is the threshold of similarity of person model, which is learning in training phase.

4 Experiments

The proposed scheme is tested by sequences taken in campus, and we named them campus sequences. In campus sequences, there are two cameras placed with no disjoint view, the topology is show in Fig.5. There are not overlapping view within two cameras. The camera handover and human tracking result is show in Fig. 6.

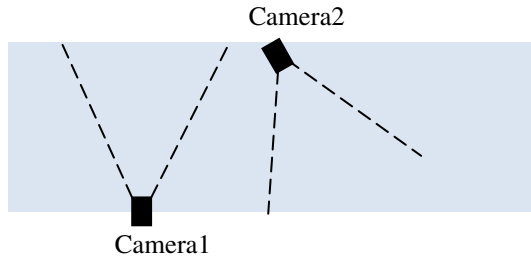


Fig. 5. The topology of cameras

Fig. 6 (a) is taken by camera1 in Fig. 5, (b) is by camera2. One person in blue shirt and black trousers is walking from left to right, after a while, he enters the FOV of camera2 (show in Fig.6 (b)). A man in gray jacket and black pants are entering the FOV of camera2 before the first person a little. Our memory-based system can identify them successfully, as show in Fig.6 (b).

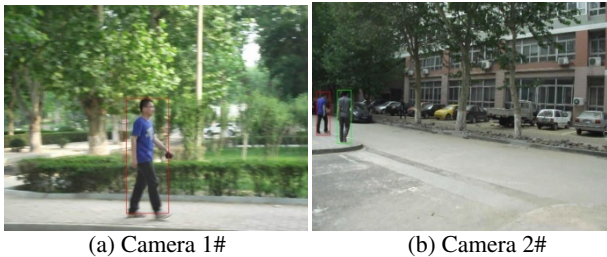


Fig. 6. Camera handover of two cameras

5 Conclusions

This paper introduced human memory mechanism to camera handover of multi-camera with non-overlapping field of view, proposed a memory-based multi-camera object tracking and handoff system. In order to reduce the computation, dominant colors of human three parts (head, torso and legs) are used as the human model for identification.

We conduct preliminary experiment use this memory-based tracking scheme and the results are satisfactory. The future research will be focused on fusion of temporal and spatial features to human model, and add other object (cars) into object list.

Acknowledgements. This work was partly supported by China Natural Science Foundation Committee (No.60975025), the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, Natural Science Foundation of Shandong Province, China (Grant ZR2011FM032), and the University Innovation Program, Jinan Science and Technology Bureau (No.201004002).

References

1. Kim, I.S., Choi, H.S., Yi, K.M., Choi, J.Y., Kong, S.G.: Intelligent Visual Surveillance - A Survey. *International Journal of Control Automation and Systems* 8, 926–939
2. Collins, R.T., Lipton, A.J., Fujiyoshi, H., Kanade, T.: Algorithms for cooperative multisensory surveillance. *Proc. of the IEEE* 89(10), 1456–1477 (2001)
3. Siebel, N.T., Maybank, S.: The advisor visual surveillance system. In: *Proc. of the ECCV Workshop on Applications of Computer Vision*, pp. 103–111 (2004)
4. Chen, K.-W., Lai, C.-C., Hung, Y.-P., Chen, C.-S.: An adaptive learning method for target tracking across multiple cameras. In: *Proc. of CVPR*, pp. 1–8 (2008)
5. Javed, O., Rasheed, Z., Alatas, O., Shah, M.: Knight: a real time surveillance system for multiple overlapping and non-overlapping cameras
6. Javed, O., Rasheed, Z., Shafique, K., Shah, M.: Tracking across multiple cameras with disjoint views. In: *ICCV* (2003)
7. Khan, S., Shah, M.: Consistent labeling of tracked objects in multiple cameras with overlapping fields of view. *IEEE Trans. PAMI* 25(10), 1355–1360 (2003)
8. Kang, J., Cohen, I., Medioni, G.: Persistent objects tracking across multiple non overlapping cameras. In: *Proceedings of the IEEE Workshop on Motion and Video Computing* (2005)
9. Kettner, V., Zabih, R.: Bayesian multi-camera surveillance. In: *CVPR* (1999)
10. Javed, O.: K, Shafique, and M. Shah.: Tracking across multiple cameras with disjoint views. In: *ICCV* (2003)
11. Toyama, K., Krumm, J., Brumitt, B., Meyers, B.: WallFlower: principle and practice of background maintenance. In: *International Conference of Computer Vision*, pp. 255–261 (1999)
12. Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: S. Sirotti. Improve shadows suppression in moving object detection with HSV color information. In: *Proceedings of IEEE Intelligent Transportation Systems Conference, Oakland, CA, USA*, pp. 334–339 (2001)
13. Li, W., Li, N., Zhang, Y.: Target tracking algorithm based on humans dominant color features in multiple cameras. *Computer & Digital Engineering* 39(4), 119–122 (2011)
14. Qi, Y.-J., Wang, Y.-J., Li, Y.-P.: Memory-based Gaussian Mixture Background Modeling. *Acta Automatica Sinica* 36(11), 1520–1526 (2010)
15. Wang, Y.X., Wang, Y.: Cognitive models of the brain. In: *Proceedings of the 1st International Conference on Cognitive Informatics, Calgary, Canada*, pp. 259–269 (2002)

The Face Recognition Algorithm Based on Curvelet Transform and CSVD

Shulin Song, Yan Zhang, Xian Wang, and Xin Mu

Key Laboratory of Advanced Process Control for Light Industry Ministry of Education (Jiangnan University), Wuxi, 214122, China
songshulin@gmail.com

Abstract. A face recognition algorithm is proposed based on Curvelet transform and Class estimated basis Space singular Value De-composition (CSVD). Face images are decomposed by using Curvelet transform firstly. As a result, Curvelet coefficients in different scales and various angels are obtained. Then, the images reconstructed by the Curvelet coefficients of the coarse layer are processed by a Fourier transform with invariant prosperity against spatial translation. CSVD algorithm is used to reduce the dimensionality and extract the feature of the amplitude spectrum face. Finally, the nearest neighbor decision rule is applied to identify the unknown face. The standard face databases of ORL, FERET and Yale are selected to evaluate the recognition accuracy of the algorithm. The results show that the proposed algorithm is used to improve the recognition rate effectively.

Keywords: face recognition, Curvelet transform, class estimated basis, Singular Value Decomposition (SVD), nearest neighbor decision rule.

1 Introduction

Face recognition is an important biometric technology, which is widely used. Compared with other identification methods, face recognition has direct, friendly and convenient characteristics. Thus, the face recognition has become a pattern recognition and artificial intelligence research focus [1]. Over the past 20 years emerged many new face recognition method, the effectiveness of these methods depends largely on the proposed extracted facial features. Currently, the features used for face recognition are the following four kinds: intuitive features, Statistical characteristics, transform coefficients characteristics, algebraic characteristics. Among them, the algebraic features of a face image reflect its intrinsic properties. SVD is an effective algebraic feature extraction method. Hong Zaiquan and others have demonstrated the stability, rotation invariance, scale invariance and shift invariance and other properties of singular value. So the singular value features are used for face recognition. Using singular value feature obtained by singular value decomposition to constitute the Sammon optimal discriminated plane. Then do recognition experiments on human's face, the result of experiments show that the singular values have a good ability to identify [2]. Later studies showed that, the singular value features obtained by singular value decomposition are inadequate used as facial image recognition features [3-4]. Tian and

others proposed that a lot of valid face information contained in the left and right orthogonal matrix by singular value decomposition of the face image matrix, and singular value vector contains a small amount of useful information of the facial image, so the recognition rate is low using singular value vector as identification feature. So the test images are projected onto the left and right orthogonal matrix obtained by SVD of the training samples. Then the projection coefficient values on the diagonal are compared with the singular value of training samples. The recognition rate is 92.5% in ORL database [3]. GAO QuanXue and others further study found that there is no one to one relationship between the singular value vector and the face image. Singular value vector is obtained in the particular base space of face image, different face images corresponding to the singular vector space where the base is inconsistent. The similarity of the base space of the same category of face image is relatively large but, the similarity of the base space of different category of face image is relatively small. Therefore they proposed a class estimated basis space singular value decomposition (CSVD) face recognition methods. The class estimated basis space is obtained by averaging the same category of face image then applying the singular value decomposition on the category average image to obtain the class estimated basis space. It is an effective way to solve the basis space inconsistent problem [4]. For the problem of basis space inconsistent, there are a lot of literature [5-7] proposed to average all the training samples as a standard image, then did singular value decomposition on it, and obtained a common base space, and finally the training samples and the test samples were projected onto the common base space. The projection coordinates is used as recognition feature. This approach solved the problem of basic space inconsistent between the samples, but ignored the category information of the samples. A new method is proposed to obtain class estimated basis space by using Curvelet transform and Fourier transform on the face image obtained amplitude spectrum image in this paper. The method first use curvelet transform to obtain sparse expression of the face, after that the discrete Fourier transform applied to obtain the amplitude spectrum image which the amplitude spectrum displacement is unchanged. Then CSVD is used to extract the recognition feature in the amplitude spectrum image. High recognition rates are obtained in main face databases.

2 Curvelet Transform

In 1999, Candes EJ and Donoho DL proposed the Curvelet concept and the theory of Curvelet transform, namely the first generation Curvelet transform[8]. But the first generation Curvelet is based on traditional structures. the concept of the curvelet is complex, the data computation is large and slow. And it is likely to cause artifact block. Therefore, Candes, made a second-generation Curvelet and fast curvelet transform. Compared with the first generation, the second generation Curvelet is easier to understand and achieve. Currently, the application of Curvelet-based research in the field of image processing has been widely carried out.

There are three parameters in the second-generation Curvelet transform, namely j is the scale, l is the direction, $k=(k_1,k_2) \in Z^2$ is spatial location. When the scale is 2^j , the rotation angle is $\theta_l = 2\pi \cdot 2^{\lfloor j/2 \rfloor} \cdot l$, $l=0,1,\dots$, Position for the translation

is $x_k^{(j,l)} = R_{\theta_l}^{-1}(k_1 \cdot 2^{-j}, k_2 \cdot 2^{-j/2})$, The defined formula of the curvelet is: $\varphi_{j,k,l}(x) = \varphi_j(R_{\theta_l}(x - x_k^{(j,l)}))$, R_{θ} is the arc of rotation on θ , $R_{\theta} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$, $R_{\theta}^{-1} = R_{\theta}^T = R_{-\theta} \cdot \varphi_j$. That is $\varphi_j(x)$, is known as the "mother wavelet". The main idea of Curvelet transform is by calculating inner product between the image f and Curvelet function φ_D to obtain the sparse representation of the image f . For digital image function $f(n_1, n_2)$, $0 \leq n_1, n_2 \leq N$, N is the dimension of images, the Curvelet coefficient is calculated as follows:

$$C^D(j, l, k) = \sum_{n_1, n_2} f(n_1, n_2) \overline{\varphi_{j,l,k}^D(n_1, n_2)} \tag{1}$$

The second generation Curvelet transform fast algorithm that is fast discrete Curvelet transform (Fast Discrete Curvelet Transforms, FDCT) [9], is divided into two algorithms: USFFT (Unequally-Spaced Fast Fourier Transform) algorithm and Wrapping algorithm. Two algorithms have the same result of the operation, but Wrapping algorithm has fast computing speed with high efficiency. So this paper uses Wrapping algorithm. The core idea of the Wrapping algorithm is wrap around the origin, through the period-one mapping technology to the origin of the affine area.



Fig. 1. Reconstructed images of Curvelet coefficient on coarse layer

Then, the reconstructed images on coarse layer Fourier transform, two-dimensional discrete Fourier transform is defined as Eqs-2 :

$$F(u, v) = \psi[f(x, y)] = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{j2\pi(ux/M + vy/N)} \tag{2}$$

From the formula $|\psi[f(x - x_0, y - y_0)]| = |F(u, v) e^{-j2\pi(ux_0/M + vy_0/N)}| = |F(u, v)|$, we can know that the displacement of face image's amplitude spectra is the same, and therefore can be used to reduce the recognition errors caused by displacement changes.

3 The Singular Value Decomposition of Face Image

The formula of singular value decomposition of Face image matrix $A_{m \times n}$ is:

$$A = USV^T = \sigma_1 u_1 v_1^T + \dots + \sigma_k u_k v_k^T = \sum_{i=1}^k \sigma_i u_i v_i^T$$

Among them, U is a $m \times m$ dimensional orthogonal matrix, V is a $n \times n$ dimensional orthogonal matrix, $S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k)$, $\sigma_1, \sigma_2, \dots, \sigma_k$ is the singular value of face image A , and it also can be seen as the projection coordinates of face image A project onto the base space in this group $(u_1 v_1^T, u_2 v_2^T, \dots, u_k v_k^T)$, u_i is the i column vector of the left orthogonal matrix U , v_j is the j column vector of the right orthogonal matrix V , $k = \text{rank}(A)$. It is shown that face images A can be a linear representation of the base matrix $u_i v_i^T$. We can also see the singular value $\sigma_1, \dots, \sigma_k$ as the weight of base matrix.

The contribution rate of the singular value to the image is expressed as $\epsilon_i = \sigma_i^2 / \sum_{i=1}^k \sigma_i^2$ ($i = 1, 2, \dots, k$), so the larger the singular value is, the corresponding base matrix of face images contribution rate is relatively large, the smaller the singular value is, the corresponding base matrix of face images contribution rate is relatively small.

Thus we can ignore the base matrix whose weight is very small, to compress the image and reduce the dimensionality of the image. We use the t maximal singular values and its corresponding left and right orthogonal matrix vectors to reconstruct the image. The reconstructed results shown in Fig. 2. :

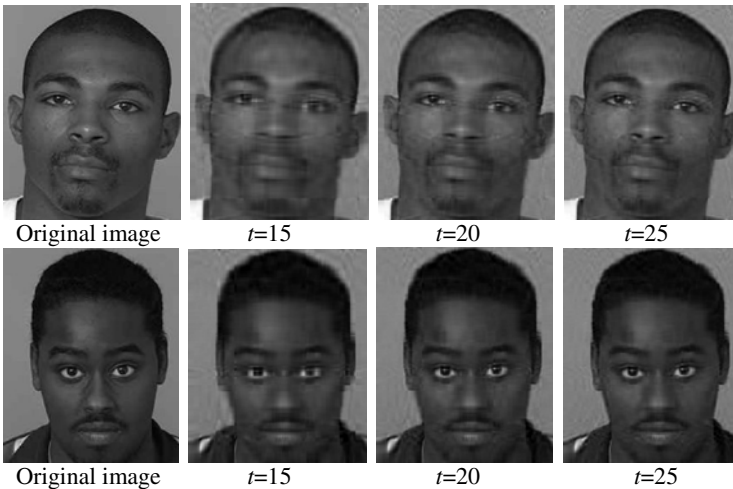


Fig. 2. Reconstructed images under different t

Fig. 2 shows, a large number of face image information can be reflected by the t maximal singular value and its corresponding left and right orthogonal matrix. Therefore, the image can be projected onto the t maximal singular value corresponding left and right orthogonal matrix; the projection coordinates S_{ij} used as the recognition features. The projection coordinates have just only $t \times t$ dimension, achieved the effect of dimensionality reduction. Since the base spaces of any two different face images are different, and comparing the size of projection coordinates in the different basis spaces is meaningless. But the same class people's basis space is similar, so we use the class estimated basis space to solve the problem of the same class people's basis space is inconsistent. And it is not ignore the information between the categories. Class estimated basis space singular value decomposition algorithm whose main idea is to find the class estimated basis space. First, the image of each class of training samples are averaged, and then do the singular value decomposition on the average image, the left and right orthogonal matrix ($U^{(i)}$ and $V^{(i)}$) are constituted the class estimated basis space of this class. Then the training samples are projected to itself class estimated basis space, $S_{ij} = U^{(i)}(:, 1:t)^T A_{ij} V^{(i)}(:, 1:t)$, obtained the projection coordinates S_{ij} , then the test image project to the each class estimated basis space to obtain projection coordinates St_i . Finally, the Euclidean distance is calculated between projection coordinates of the test image in the i class estimated basis space St_i and projection coordinates of i class training sample in its estimated basis space coordinates S_{ij} , and then find out the minimum to represent the test image with the similarity of the i class image, according to the minimum of similarity value you can determine the category facial image to be identified belongs to.

4 The Experimental Procedure and Experiment Result Analysis

4.1 Experimental Procedure

(1) Apply the scale=3 and angle=8 Curvelet transform on the image A_{ij} , then apply the Fourier transform on the reconstructed image of Curvelet coefficients on the coarse layer then obtain the amplitude spectrum images A_{ij}' .

(2) Compute the i category average amplitude spectrum of training sample set $A_i' = \sum_{j=1}^{Nc} A_{ij}'$ ($i = 1, 2, \dots, c, j = 1, 2, \dots, Nc$) i represents the number of categories of training samples, j represents the number of training samples of each category.

(3) Apply the singular value decomposition on the A_i' , take the t maximal singular value corresponded the left and right orthogonal matrix vector u_p, v_p ($p=1, 2, \dots, t$) as the i class estimated basis space. So the i class estimated basis space are $U_i = [u_1, u_2 \dots u_p]$ and $V_i = [v_1, v_2 \dots v_p]$ ($i = 1, 2, \dots, c$)

(4) Project the amplitude spectrum of training samples onto the each class estimated basis space, $S_{ij} = U_i^T A_{ij}' V_i$ ($i = 1, 2, \dots, c, j = 1, 2, \dots, Nc$) i represents the number of categories of the training face images, j represents the number of training samples

of each category. S_{ij} represents the projection coordinates of the i class, the j th face image in the i class estimated basis space.

(5) Apply the Curvelet transform and Fourier transform on the test image to obtain the amplitude spectrum image I , then compute the projection coordinates in each class estimated basis space $St_i = U_i^T IV_i$.

(6) According to the nearest neighbor decision rules to classify, first, compute the minimum Euclidean distance between the projection coordinates of the test image in the i class estimated basis space and the projection coordinates of the i class training samples in itself class estimated basis space. And regard the minimum Euclidean distance as the similarity of the test image and the i class image. Then find the minimum similarity, the category that the minimum similarity corresponding to is the recognition image belongs to.

4.2 Experimental Results and Analysis

Testing the proposed algorithm on the ORL FERET Yale JAFFE face database, The ORL face database is created by the University of Cambridge; there are 40 people in the database, 10 photos of each people, including small changes in posture, 20 degrees within the scale changes. Take 5 photos as training samples, remaining as the test samples, so there are 200 photos in the training samples and test samples. The FERET face database created by the FERET program, including more than 14,051 postures, light gray face image, in which we selected 20 individuals, each person selected 10 pictures, 5 of them for training and the other five for testing. The Yale face database created by the Yale University, there are 15 people, 11 photos of each people, including changes in light conditions, and changes in expression, Take five photos of each people for training and another six for testing. The Japan JAFFE facial expression database includes 216 facial expression images. 20 individuals, 10 photos of each, are selected in the experiment. The results are shown in Figure 3.

It can be seen from Figure 3, the recognition rate can be as high as 99.25% in the ORL face database when $t = 30$, but the features are extracted 30×30 dimension; when $t = 25$, the recognition rate is also 98.9%, and at this time, extracted features only 25×25 dimension, it reduced low computation and less time. The highest recognition rate can be 94.43% in the Yale face database. Recognition rate in the FERET face database can also be 88.11%.

Take $t = 25$ to compare recognition rate of the algorithm in the ORL, FERET, Yale, JAFFE face database with the different number of training samples. Results are shown in Figure 4:

It can be seen from Figure 4 that, even in the case of only one training samples, the recognition rate of the algorithm in the ORL and FERET face database achieve to 75.28% and 76.67% respectively. The recognition rate in the Yale face database is 45.33%.

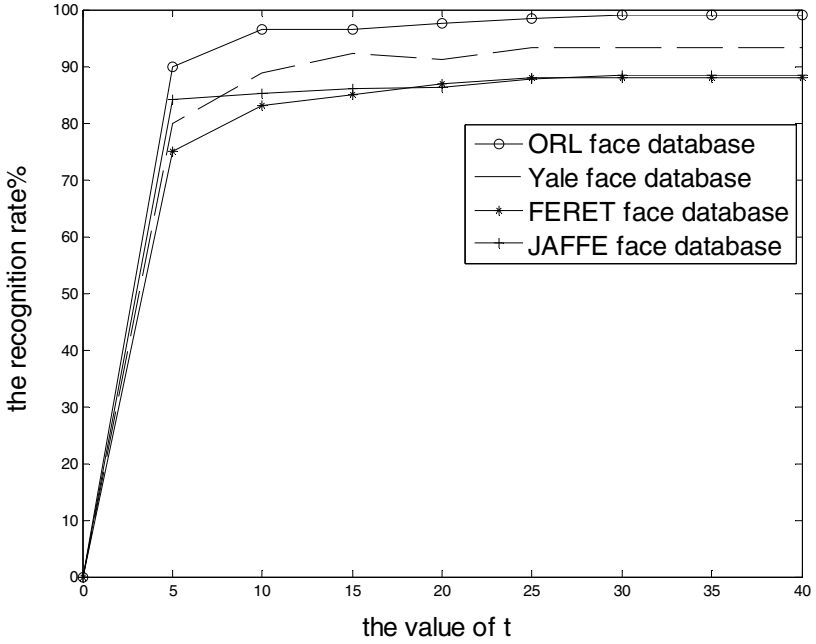


Fig. 3. Recognition rate according to different t of this paper algorithm

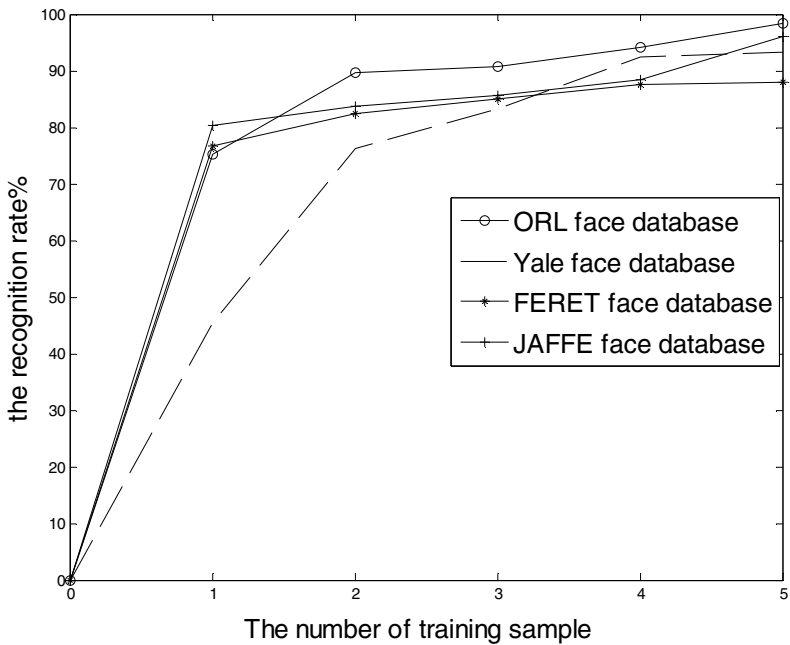


Fig. 4. Recognition rate according to different sample number of this paper algorithm

Using CSVD algorithm (method 1), mean face SVD[5] (method 2), wavelet transform + mean face SVD (method 3) and the proposed algorithm, respectively, in ORL, YALE, FERET, JAFFE face database for comparison, the comparison results shown in Table 1:

Table 1. Recognition rate of different algorithm in the ORL database

training sample number	method 1	method 2	method 3	this paper
3	87.8%	88.2%	89.6%	90.7%
4	90.8%	91.6%	92.7%	94.1%
5	93.2%	94.7%	97.1%	98.9%

Table 2. Recognition rate of different algorithm in the YALE database

training sample number	method 1	method 2	method 3	this paper
3	80.9%	81.7%	82.9%	83.3%
4	88.6%	90.5%	91.6%	92.3%
5	91.4%	91.9%	93.8%	94.4%

Table 3. Recognition rate of different algorithm in the FERET database

training sample number	method 1	method 2	method 3	this paper
3	83.3%	83.8%	84.5%	85.2%
4	84.1%	85.2%	86.9%	87.1%
5	85.6%	86.5%	87.4%	88.1%

Table 4. Recognition rate of different algorithm in the JAFFE database

training sample number	method 1	method 2	method 3	this paper
3	83.4%	84.1%	85.1%	85.7%
4	85.2%	85.9%	87.6%	88.3%
5	90.0%	90.3%	95.2%	96.5%

5 Conclusion

A face recognition algorithm is proposed based on Curvelet transform and CSVD, which used the Curvelet transform to obtain the sparse expression of face and dimension reduction, effectively reduce the computation time. Then the CSVD algorithm is used for feature extraction, which is benefit to the basis space consistence, not ignoring the sample information between the categories. In future the left and right

orthogonal matrix of to the singular value decomposition will be considered to obtain better estimates of class-based space.

References

1. Duan, J.: Face automatic machine recognition. Science Press, Beijing (2008)
2. Hong, Z.-Q., Yang, J.-Y.: The image algebra feature extraction used in image recognition. *Automatic Acta* 18(2), 233–238 (1992)
3. Tian, Y., Tan, T.-N., Wang, Y.-H.: Do singular values contain adequate information for face recognition. *Pattern Recognition* 36(6), 649–655 (2003)
4. Gao, Q.-X., Lian, Y., Pan, Q.: The Existed Problems of SVD in Face Image Recognition and Related Solutions. *Chinese Image Graphics Journal* 11(12), 1785–1791 (2006)
5. Hsu, C.-H., Chen, C.-C.: SVD-Based projection for face recognition. In: *International Conference on Electrol Information Technology*, pp. 600–603 (2007)
6. Guo, Z.-Q., Yang, J.: Method to extract features of face image combined SVD and LDA. *Computer Engineering and Design* 30(22), 5133–5139 (2009)
7. Guo, Z.-Q., Yang, J.: A new method to extract face features based on combination of mean face SVD and KFDDA. In: *International Conference on Biomedical Engineering and Computer Science*, pp. 1–4 (2010)
8. Candes E. J, Donoho D. L.: Curvelet-a surprisingly effective non-adaptive representation for objects with edges. *Curves and Surfaces*, 105-120 (2000)
9. Emmanuel, C., Laurent, D., David, D.: Fast Discrete Curvelet Transforms. *J. Applied and Computational Mathematics*, 1–43 (2005)

Background Modeling for Fire Detection

Yan Yang^{1,2}, Xiaopeng Hu¹, Chaofei Zhang¹, and Yi Sun¹

¹ School of Information & Communication Engineering, Dalian University of Technology,
Dalian 116012, China

² School of Computer Science & Information Technology, Liaoning Normal University,
Dalian 116029, China

Jasmineyy@mail.dlut.edu.cn, {xphu, lslwf}@dlut.edu.cn,
workroomofzcf@126.com

Abstract. A video-based fire detection system for outdoor surveillance should be capable of continuous operation under various environmental conditions, such as the change of background illumination and variation of background objects. In this paper, we present a novel background modeling method in which a nonparametric statistical test is utilized for effective and robust detection of stationary background corners. The practical value of the method is demonstrated with an oil-field surveillance system where it is applied for video-based fire detection. The validation results and analysis indicate that the proposed method is able to cope with small occlusions and periodic motions.

Keywords: background modeling, Harris corner, nonparametric test, fire detection.

1 Introduction

Background modeling for robust motion detection is a fundamental problem for foreground and background extraction in video surveillance applications. In a video-based outdoor fire detection system, for instance, the background modeling, which is the focus of this paper, is the basis for fire detection by determining which pixels or regions in a video frame fall into the foreground or background classes.

Most existing methods [1-5] operate in the spatial domain at pixel levels. Gaussian models are widely applied for the representation of scene background [1-2]. In complicated natural scenes, Gaussians mixture modeling (GMM) [2] is proposed to handle illumination changes and background-object variations since each independent pixel can't be properly modeled by a single Gaussian distribution[1]. To make less assumption about the distribution of background pixels, Ahmed Elgammal[3] proposed a nonparametric kernel density method to support robust background modeling in dynamic outdoor scenes with small motions such as swaying of grasses, trees and their shadows. To decrease the computational complexity, Qiang Zhu[4] proposed a corner-based model based on a set of SIFT-like features. As the number of detected corners can be controlled by adjusting the parameters, such sparse background representation is computationally more efficient in the aspect of constructing and maintaining the model. The drawback of the model is improved by Muhammad Uzair[5]. Both

the algorithms apply a 128-dimension feature vector to model every background corner. However, such a high dimensional feature space is memory consuming.

This paper proposed a novel robust and sparse background modeling method for fire detection in a video-based oil-field surveillance system where accuracy, robustness, computation and memory requirements are important factors. The video stream is obtained by static cameras. The method utilizes color features [6] to detect regions of interest (ROIs) where the Harris corners are calculated. A nonparametric statistical test is applied for robust detection of background corners and an updating scheme of the background model is designed to ensure the tolerance of temporary occlusions.

2 ROIs Detection and Sparse Corner Extraction

It is found that the geometric structure of fire flames can be represented by corner detectors. In this paper, standard Harris corners[7] are detected in ROIs for background modeling. Harris corner detector is robust to image-gradient variations, such as the illumination changes. The image gradients I_x and I_y are calculated to measure the intensity variation in principal directions. The cornerness value $R(x,y)$ at location (x,y) is calculated as follows:

$$R(x,y) = \text{Det}(M) - k \times \text{Tr}^2(M) . \quad (1)$$

where k is a constant which is usually set between 0.04 and 0.06 [7], and Det and Tr are the determinant and trace of autocorrelation matrix M respectively. Given a pixel $I(x,y)$, it is considered as a corner point if $R(x,y) > \text{thre}R$.

In a complicated outdoor oil-field scene, the corners will be misclassified as foreground corners and it will increase the difficulty of fire detection. To avoid the unnecessary effect from them and to decrease computational complexity, ROI extraction is performed in this paper. In [6], P.V.K. Borges proposed a fire detection method for hydrocarbon flames using color clues to indicate candidate fire regions since color is the most discerning feature. Such method is based on the assumption that all the distributions of red, green and blue channels follow Gaussian distribution. However, it's not true for the near infrared cameras of HIKVISION whose type is DS-2CD802P(N)-IR3. With the default parameters of the camera, flame color falls into the white range instead of red-yellow range[6]. Fig.1 shows the RGB histogram of a fire region inside the red square. Most fire color intensities fall into the range between 230 and 255. Based on this fact, we simplify the ROI extraction procedure and use the same relaxed threshold $\text{thre}I$ to all the color channels to indicate the fire candidate regions. The swaying green vegetation and their shadows can be easily classified as background since their color is out of the region-of-interest color range.

3 Background Modeling Based on Nonparametric Statistical Test

In practice, it is difficult to identify the distributions of a background corner along with certain time when moving objects passing through the background pixel or background illumination changing because of the weather. This paper proposed a novel background modeling framework based on sign test. It is not necessary to as-

sume any particular shape for the distribution of the background corner.

Sign test [8] is one of the simplest nonparametric statistical tests. It works on the theory that there are equal probabilities of two outcomes. For the same background corner in frame sequences, position differences between every two consecutive frames are calculated and its signs are recorded as 1 and -1 . If there is no corner detected in the corresponding position of the previous frame, the difference will be calculated between the latest frame which has the corresponding corner and the current frame. The hypothesis of sign test is expressed as follows:

$$\begin{aligned} H_0: p &= 0.5 && \text{It is background corner .} \\ H_1: p &\neq 0.5 && \text{It is foreground corner .} \end{aligned} \quad (2)$$

where p is the proportion of $+$ signs or $-$ signs. If the null hypothesis H_0 is true, the corresponding corner is background corner. Otherwise, it is marked as foreground corner.

The proposed background modeling algorithm is as follows:

- 1) For every new frame, detect Harris corners inside ROIs;
- 2) The color metrics Dis_{RGB} is calculated using Euclidean distance[12] to associate the new corners with existing background corners and the neighborhood size is 5×5 ;
 - If Dis_{RGB} is greater than the threshold $thre_{RGB}$, the new corner will be classified as a foreground corner. Test1 is performed on such unassociated corner to detect new background corner.
 - Otherwise, mark the associated corner as background corner and perform Test2 to check all other background corners that are not associated to any new corner.
- 3) Update the background model according to the associated new corners.

In particular to mention, sign test can improve the robustness to small periodic or repetitive motions, such as jittering sensors and vegetation swaying in the wind, because the sign test only focuses on whether there are equal probabilities of the relative position variations or not. Test1 is to apply the sign test on frame sequences containing the latest $T1$ successive frames. Only non-background corners are considered. Within the sequence, the maximum number of successive frames that don't contain corners should be less than $L1$. When a moving object, such as a parking car, becomes stationary over a long period of time, new corners that belong to the moving object will be incorporated into the background model and old background corners which are occluded by the moving object will be removed from the background model. Test2 is to apply the sign test on frame sequences containing $T2$ successive frames within the latest $T3$ frames ($T3 > T2$, set $T1 > T2$ to overcome occlusions). Only background corners are considered. Within the sequences, the maximum number of successive frames that don't contain corners should be less than $L2$. If the parameters are properly set in accordance with real-world requirement, the occluded background corners over a short period of time will not be removed from the background model. The smaller the value of $T2$ is, the slower the speed of removing old background corners is and the higher robustness to the occlusion can be obtained. Moreover, the values of $L1$ and $L2$ determine the occlusion duration.

4 Experimental Results

The proposed method is evaluated on four video sequences of various environmental conditions. The important parameters of the proposed framework are fixed as $threI=200$, $threRGB=5$, $k=0.04$, $T1=30$, $T2=6$, $T3=20$, $L1=20$, $L2=6$ through the experiments.

In this paper, both visual and numerical methods are used to evaluate the proposed approach. Fig.2 shows two results of different outdoor scenes with both illumination changes and background moving objects. The first row shows four frames from the practical video sequence. The images in the second row show the candidate fire regions with foreground corners marked as red \times . In the third row, outdoor scenes with both background and foreground corners are given. To make a numerical comparison, 30 frames are randomly selected from each sequence (i.e. 200 frames in total) and manually labeled as ground truth. Table.1 illustrates the corner-level false negative (foreground corners that are misclassified as background) rate and false positive (background corners that are incorrectly marked as foreground) rate produced by our approach. The false negative rate is less than 0.128 % and it validates the robustness of the sign test. The false positive rate is a little higher, above 0.49 %, but it is believed that it could be reduced with further fire feature analysis.

Table 1. The false negative and the false positive rate of the proposed approach

Seq No.	False negative	False positive
1 st	0.103%	0.86%
2 nd	0.114%	0.78%
3 rd	0.119%	0.57%
4 th	0.128%	0.49%

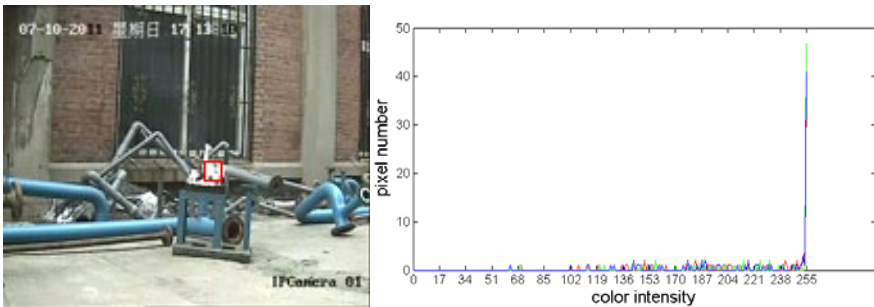
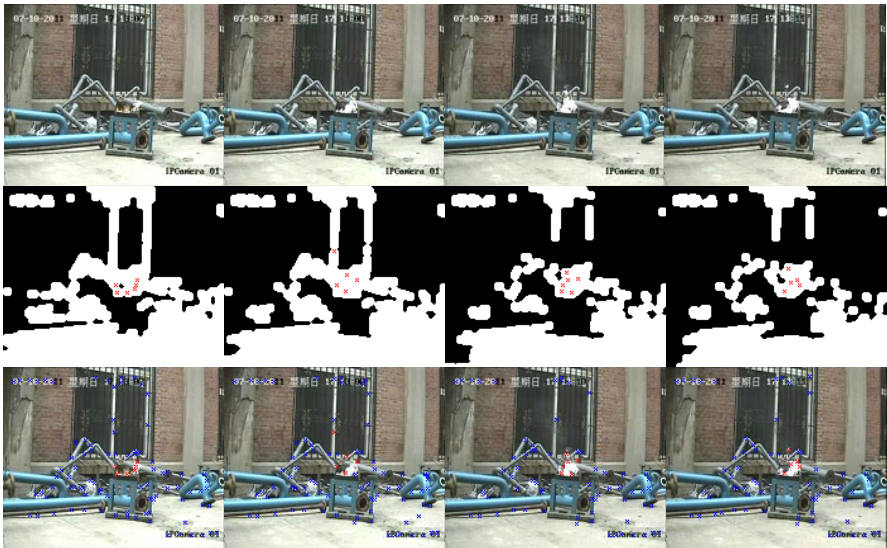


Fig. 1. RGB Histogram of a fire region inside the red square

We also measured the speed of the proposed method. The Matlab code was conducted on a computer with an Intel 2.4GHz i5 CPU and 2GB memory. For the parameter values used in the tests, a frame rate of 18 fps was achieved in average and the resolution of the test video is 176×144 . The memory consumption can be saved about 20% compared to [4].



(a) 2nd Video(500f,575f,650f,725f out of 1800 video frames in total)



(b) 4th Video(350f,400f,450f,500f out of 1000 video frames in total)

Fig. 2. Videos with illumination changes and grasses swaying in the wind

5 Conclusions

The purpose of this paper is to establish a corner-based approach for reliable background modeling in complicated natural environment where variations and background object movements are widely presented. The proposed framework applies a nonparametric statistical test that operates on sparse corners inside ROIs to build a background model. The dissimilarity of colors is utilized to differentiate background corners from foreground corners. A background updating scheme is designed to overcome temporary occlusions. Experimental results and analysis demonstrate the robustness and computational efficiency of the proposed framework, which has been applied in a practical oil-field surveillance system.

Acknowledgments. This research was supported by “National Natural Science Foundation of China” (No. 61075018), “the National Key Project of Science and Technology of China” (No. 2011ZX05039-003-4) and “the Fundamental Research Funds for the Central Universities”.

References

1. Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.: Pfnder: Real-Time Tracking of the Human Body. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19 (1997)
2. Stauffer, C., Grimson, W.E.L.: Adaptive Background Mixture Models for Real-Time Tracking. In: *CVPR*, vol. 2, p. 2246 (1999)
3. Elgammal, A., Harwood, D., Davis, L.: Non-parametric Model for Background Subtraction. In: Vernon, D. (ed.) *ECCV 2000*. LNCS, vol. 1843, pp. 751–767. Springer, Heidelberg (2000)
4. Zhu, Q., Avidan, S., Cheng, K.-T.: Learning a Sparse Corner-Based Representation for Time Varying Background Modelling. In: *ICCV*, vol. 1, pp. 678–685 (2005)
5. Uzair, M., Khan, W., Rehman, F.: Background Modeling Using Corner Features: an effective approach. In: *IEEE International Multitopic Conference on INMIC*, Islamabad, Pakistan (2009)
6. Borges, P.V.K., Izquierdo, E.: A Probabilistic Approach for Vision-Based Fire Detection in Videos. *IEEE Trans. Circuits Syst. Video Techn.* 20, 721–731 (2010)
7. Harris, C., Stephens, M.J.: A Combined Corner and Edge Detector. In: *Proc. Alvey Vis. Conf.*, pp. 147–152 (1988)
8. Howell, D.: *Psychology for Statistics*. Duxbury Press (1997)
9. Cha, S.: Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions. *Int. J. Math. Models Meth. Appl. Sci.* 1, 300–307 (2007)

Research on the Technology of Video Semantic Retrieval Based on Structured Semantic Strings

Jing Yuan, Quan Zheng, Zhijun Sun, and Song Wang

Department of Automation, University of Science and Technology of China,
Hefei, 230027, P.R. China
phoenix8@mail.ustc.edu.cn, qzheng@ustc.edu.cn

Abstract. To explore a concise and efficient approach to address video semantic retrieval problem, we creatively propose a concept of structured semantic string to realize video indexing and retrieval in this paper. Integrated with the technology of NLP (Natural Language Processing) and documents inverted index, our method could help represent information on sentence level so that computer could better understand videos' descriptions. In addition, expansion on synonyms and hypernyms contributes to comprehensive semantic expressing. Presently, the approach has been applied in our distributed video semantic retrieval system – Xunet, and it is proved to achieve high recall rate and precision rate.

Keywords: video semantic retrieval, NLP, semantic expanding.

1 Introduction

Facing a bunch of videos offered by video service websites, users depend on video search engine the same as they rely on general web search engine. How to provide videos in demand precisely is a hot problem puzzling both industry engineers and academic researchers. Apparently, the vital point to increase the retrieval effect is representing video information and search intensions semantically.

Nowadays, mainstream video retrieval technologies still rely on keywords matching. Text annotations attaching to videos help describe video information. These contents are then converted to words frequency vectors and stored in inverted indexes documents. Keywords refined from search contents are submitted to be retrieved in index [1]. Though the technology has been maturely utilized, it is not enough to accommodate users' demands any more. First, over-reliance on term frequency in a document to judge the similarity between a document and a term is not reasonable considering the interference of massive ineffective words repetitively appear in a document. Second, owing to lacking in analyzing relationships of words in a sentence, computer could only understand descriptions on words level rather than sentence level. That is to say, it may break up the comprehensive meaning of videos.

From the perspective of videos' low-level features analyzing, CBVQ (content-based visual query) technology which attracts many experts' attentions. Researchers are trying to depict videos through still key frame's characteristics such as color, texture and shape or motion trajectory in time domain like object tracking [2, 3]. There

has already been some accomplished system on this theme. For instance, VideoQ of Columbia University [2], BilVideo of Bilkent University [3] and Multimedia Search and Retrieval System of IBM [4]. Nevertheless, it is difficult to use image's low-level features to reflect high-level semantic information. Besides, dimension disaster prohibits this idea smoothly applied in practical systems. As to object motion trajectory method, object shape recognition hasn't been perfectly solved yet. Therefore, many attempts on CBVQ solely focus on a specific field such as sports [5].

Ontology is a concept thriving in recent years and contributes to specific field to establish its structured and standard description. Admittedly, as a prospective method to describe a small world, ontology acquires W3C's adequate attention so that its standards such as RDF (Resource Description Framework) and OWL (Web Ontology Language) are formulated in a few years. These standards provide general terms for describing knowledge in domains [6]. Some researchers are trying to apply the concept of ontology in video information indexing and retrieval [7]. Unfortunately, domain dependency restricts ontology researching range to a tiny circle. Unlike a specified field such as pharmacy and marketing chain that possess their own criteria to refer to, it is improper to convert various video descriptions to a limited standard form deficient of vivid semantic information.

Just as drawbacks mentioned above, common users are fastidious with searching effects provided by common popular technology while novel methods are under research and many obstacles are intercepting a giant leap. In this paper, we maximize the usability of matured technology and creatively propose the concept of structured string to semantically present video information. We utilize the technology of NLP (Natural Language Processing) to process video description and generate them to structured semantic strings automatically. Inspired by the concept of ontology and a semantic image annotation tool Caliph, we divide processing results to several levels' strings. These steps could filter much of useless words or information and express sentences as 0, 1, 2 level of strings respectively. Afterward, synonyms and hypernym expansions of searching contents further construct comprehensive understand of searching intentions. If only descriptions comprise proper sentences, their meanings could be stored in inverted documents concisely and semantically.

The rest of this paper is organized as follows, section 2 introduces some related technologies and works, section 3 describes the semantic expression procedure, section 4 exhibits the experimental results of our approach, section 5 makes a short conclusion to the work we have done.

2 Related Works

It is undeniable that our work has borrowed some ideas from NLP, inverted indexes and Caliph's scene description, yet to our best knowledge, it is the first time to aggregate these elements in a Chinese video retrieval system to overcome semantic gap problem. These significant related technologies and works will be introduced in this section.

2.1 Caliph and Emir

Caliph and Emir are MPEG-7 based Java prototypes for digital photo and image annotation and retrieval. They support graph like annotation for semantic metadata and content based image retrieval using MPEG-7 descriptors [8]. Caliph offers a convenient user interface for users to annotate images manually. It describes a scene by a semantic graph which is consisted of entity nodes and relationships between these nodes. Emir could then retrieve these annotated MPEG-7 files in database or Lucene index. However, Caliph stipulates so many semantic relations that users' personal preference will cause different annotation results about a same image and it will definitely affect annotating precision and retrieval effect.

In our video retrieval system Xunet, we borrow the idea from Caliph to create our structured semantic strings. We also use the concept of 0, 1, 2 levels of strings to depict a scene, 0 level string represents a single entity, 1 level string contains two nodes and their relationship, 2 level string involves three nodes that represent subject, property and object respectively. These strings that are generated by NLP could explain a specific scene appropriately and could be treated as concrete terms when build index. What's more, that automatic process makes it possible for users to use natural language to describe a scene or an event averse the troublesome brought from manual annotation. Table 1 below shows an example of 3 different levels of structured semantic strings converted from a sentence. All sentences are translated from Chinese.

Table 1. Example of structured semantic strings

Form	Components
Sentence	I loveBeijing Tian'anmen Square.
0-Level Strings	{ I, love, Beijing-Tian'anmen-Square }
1-Level Strings	{ AgentOf_I_love, ObjectOf_love_Beijing-Tian'anmen-Square }
2-Level Strings	{ SVO_I_love_Beijing-Tian'anmen-Square }

2.2 Chinese Natural Language Processing

Unlike English which uses backspace as splitter between words, Chinese sentence possess no tag to distinguish different word. Both Chinese character and characters combination could represent a word. This feature results in the reliance on analyzer to split sentences automatically. In this paper, we make use of Chinese Analyzer which is created by Chinese Academy of Sciences to complete words splitting. Chinese Analyzer is a Java encapsulated tool package, which refers to related lexical library.

LTP (Language Technology Platform) is a language processing system framework designed by Harbin Institute of Technology. It uses XML-based texts to exhibit processing results and offers an effective bottom-up Chinese processing module, including some core technology as lexical analyzing and grammar recognition [9]. With the help of LTP's DLL application program interface and dependency tree, we add more functions such as sentences filtering and qualifiers extracting to form our own language processing procedure to cater for the demand of video description.

2.3 Lucene: Multiple Fields Retrieval

Lucene is an inverted full-text search engine tool which was originally developed by Doug Cutting and currently it is a subproject of the Apache Software Foundation project team. Lucene treats a string as a term and defines document consisted of many strings, thus a document could be represented as a vector of terms. Elements of these vectors are frequencies of these terms. Through vectors calculation, similarity between term and document could be computed by a specific formula defined in Lucene.

Moreover, Lucene provides practical interfaces to realize multiple fields' retrieval so that users could retrieve documents by their interesting theme. For example, we could separate description content to several themes and build up indices respectively. In this paper, we store keywords and structured semantic strings in different fields' index to compare their retrieval effects. Currently, we have realized three methods of retrieval: keywords retrieval, semantic graph retrieval and semantic structured string retrieval. To contrast their performances, we establish searching fields for each of them. We also provide three user interfaces to test retrieval effects.

2.4 Xunet Video Semantic Retrieval System

Our project named Xunet is a distributed video semantic retrieval system mainly consisted of video storage subsystem, video preprocessing subsystem, video retrieval subsystem and semantic processing subsystem. The architecture of this system is shown in Fig. 1. Reference [10] introduces the detailed system design and implementation, including deploying procedures. The subsystem in red rectangle in Fig. 1 is the major point we focus on in this paper.

As can be seen from Fig. 1, semantic processing subsystem, distributed indexing and retrieval subsystem are vital parts of this project which could directly determine the users' experience and evaluation. In the next section, we will emphasize on the semantic processing subsystem module and demonstrate its devotion to video semantic retrieval effect.

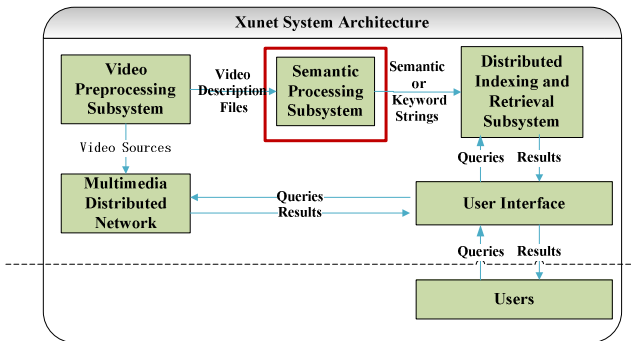


Fig. 1. Xunet system architecture

3 Video Semantic Retrieval

In this section, we will focus on detailed procedures of semantic conversion of video description. How to extract useful semantic segments which reflect video description's characteristic precisely? How to distinguish weights of different levels of semantic strings? What is a proper similarity calculating method? All these questions will be answered as following.

3.1 Semantic Structured Strings Formation

1) Semantic String Generating

To acquire structured semantic strings, we use semantic processing module to deal with it automatically. The process of this module is exhibited in Fig. 2 and described as follows.

a) Grammar Process Sub-module: Make use of LTP, mainly complete tasks of words splitting, part of speech recognition and determining grammatical relations between mutually dependent words (SBV, ATT, VOB etc.). *b) Sentence filtering sub-module:* Help judge and discard some useless or unconvertible sentences. *c) Main sentence extracting sub-module:* Remove some adverbials and attributes based on dependent relations left behind. *d) Sentence pattern recognition sub-module:* Analyze components of main sentence and determine the pattern of sentence. Mainly emphasize on five sentence patterns: declarative sentence, linked sentence, inversions, predicative sentence and complex sentence. Chinese sentence patterns are far more than these, while these five kinds are sufficient to basically represent video information. *e) Qualifiers extracting sub-module:* Conserve some potential useful qualifiers such as attributes and complements. *f) Semantic strings generating sub-module:* Express previous parsing results as structured semantic strings.

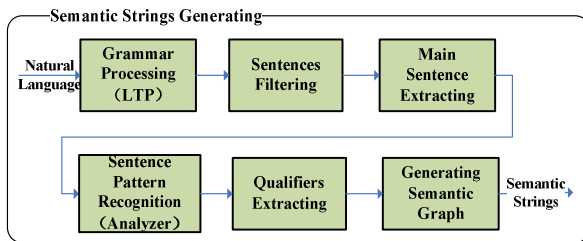


Fig. 2. Structured semantic strings generating

2) Relationship Selection

Among all three levels of generated semantic strings, it is evidently that 1-level strings and 2-level strings could better represent a sentence's integrity meaning. Since relations that connect entity nodes decide strings' form and structure, we have carefully chosen several relations which could not only express general videos' information but also simplify processing steps. Generally speaking, video content is consisted of scenes and events. Inferred from this opinion, we could draw the idea that entity

nodes could make up scenes and entity nodes accompanied with relations could represent specific events. If so, timing, location and agent or object could be treated as main elements of a video. As a matter of fact, according to the statistics of video descriptions we have acquired from video service websites by crawlers, these four elements indeed outweigh other elements. Therefore, we utilize these four relations (TimeOf_, LocationOf_, AgentOf_, ObjectOf_) to connect entity nodes. Experimental results in section 4 could demonstrate its reasonability.

3.2 Semantic Expansion

1) Single Node Expansion Procedure

Merely representing a sentence as a structured semantic string regardless of the existence of synonyms and relevant hypernyms may omit some familiar results when retrieve information among video descriptions. Only when all these relevant contents are taken into consideration, may the research intention be clearly represented. Hence we expand the searching content scope to synonyms and hypernyms. Our synonyms library derives from Chinese Synonyms Forests while hypernyms mapping table comes from Chinese WordNet, which was originally innovated by University of Princeton in English version and was than translated by Southeast University to Chinese version. Fig. 3 illustrates these two expansion procedure.

Each word has a serial number in Chinese Synonym Forest and a location number in WordNet. Though these two knowledge libraries own different encoding method, the function of their specific serial numbers or locations is to make similarity calculation more directly. There is always a minimum threshold of similarity or minimum level for expansion. Previous to the two expanded sets coming out and merging, words in the set should have already been filtered to meet threshold requirement so that irrelevant words won't infect final results.

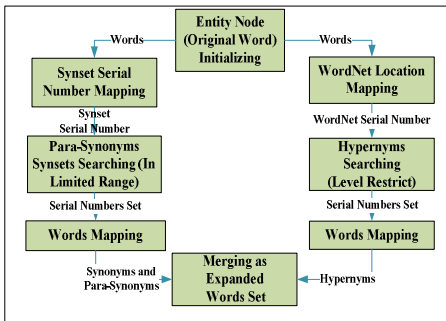


Fig. 3. Words expansion data flow diagram

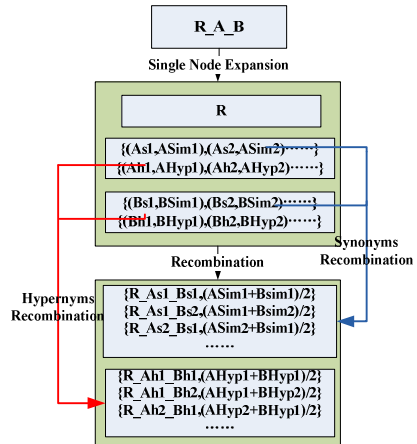


Fig. 4. Semantic strings' recombination

2) Structured Semantic Strings' Recombination

As the procedure introduced above, it only shows the expansion of a single entity nodes and the condition is only suitable for 0-level strings. Whereas for 1-level and 2-level semantic strings, after semantic expansion, more work should be done to recombine these expanded words as an entirety.

Take a 1-level semantic string's expansion recombination as example. Fig. 4 illustrates the detailed formation. Here we define R_{A_B} as a specific structured semantic string generated by natural language processing module mentioned above. In this figure, R means the relationship of entity node A and entity node B . Naming principle of A 's relevant words is that $As1$ indicates one of A 's synonyms, $Ah1$ indicates one of A 's hypernyms, $Asim1$ means the similarity between $As1$ and A , $AHyp1$ means the hypernym level degree between $Ah1$ and A , and naming principle of B 's relevant words could be inferred then. Red arrows sign A and B 's hypernyms recombination, blue arrows sign A and B 's synonyms recombination.

When word A and word B go through the expansion process, they could be transformed to two sets, one is synonyms set, the other is hypernyms set. These sets are comprised of pairs consisted of relevant strings and similarity or level degrees. When recombination completed, the fresh semantic string will own its specific weight value. Synonyms combination's weight equals to half of the sum of $ASim$ and $BSim$, hypernyms combination's weight equals to half of the sum of $AHyp$ and $BHyp$. Consequently, the expanded set is made of pairs of structured semantic string and its proprietary weight.

3.3 Weights Setting and Similarity Calculation

1) Level Weights Setting

Due to our retrieval mechanism based on inverted index, strings frequency in a document still determines retrieval results' matching degree. So it is unavoidable that 0-level semantic strings possess the largest probability to be retrieved. However, only 1-level strings and 2-level strings could evidently better exhibit a sentence's whole meaning. Large amount of 0-level strings involved in description will definitely spoil the retrieval accuracy. Take an example of a sentence which means "Clouds are floating." in English. It is possible that documents with high frequency of 0-level string "Clouds" or "floating" would acquire higher similarities and ranks while those documents containing "Cloud is floating" with one appearance are easily be omitted or ranked lower. With regard to this problem, we stipulate specific weight for each level's strings to enhance the importance of high level strings. Current level weights setting condition is listed in Table 2.

Table 2. Level weights setting

String's Level	Weights Parameter	Value
0-Level Strings	ZERO_CLASS	0.2
1-Level Strings	FIRST_CLASS	1.0
2-Level Strings	SECOND_CLASS	5.0

2) Inner Words Weights Setting

In the procedure of grammar analyzing, to maximize the information retaining, we sometimes bound adverbs or state qualifiers with nouns as an entity. For instance, in a structured semantic string which means “SVO_I_love_Beijing-Tian’anmen-Square” in English, “Beijing-Tian’anmen-Square” is a location entity and three words in this semantic string all contain important information. It is necessary to disassemble the original entity to an entity set. In this instance, original string will be transformed to a set consisted of three strings: “SVO_I_love_Beijing”, “SVO_I_love_Tian’anmen” and “SVO_I_love_square”. Thus location information could be separated and dispersed into several strings with inner words. To distinguish the original string and string with inner words, inner words weight should be considered. Table 3 shows the inner words weights setting. Such semantic strings with inner words should be multiplied by inner words weight after they multiplied by level weight.

Table 3. Inner words weights setting

String’s Level	Weights Parameter	Value
Original String	ME_SIM	1
String With Inner Words	ME_INNER_SIM	0.2

3) Synonyms’ Similarity Calculation

In order to discriminate synonyms and original words’ matching degree, we use similarity calculation to determine the weight of a semantic string after synonyms expanding. Fig. 5 expresses synonyms forest structure and Chinese words’ encoding formation.

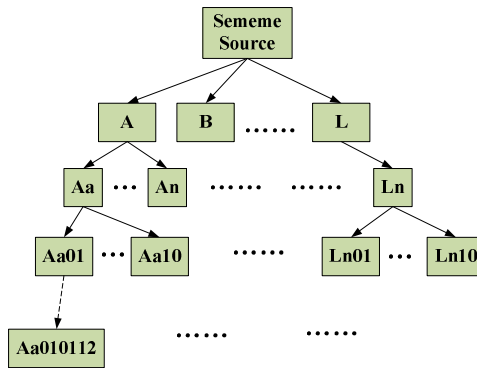


Fig. 5. Synonyms forest structure

Distance between words is a key standard to measure the relevance of two words. In many cases, directly computing the similarity is difficult so that distance between two words is always calculated at first step, and then it could be transformed to words’ similarity [11]. Words distance is a real number in the range of $[0-\infty]$ and the distance between a word and its own is 0. In addition, there exists a principle, the greater the distance is, the lower the similarity will be.

As for two words w_1 and w_2 , we express the distance between them as $Dis(w_1, w_2)$, and the similarity as $Sim(w_1, w_2)$. Then we define a simple transformation meeting above conditions:

$$Sim(w_1, w_2) = \frac{\alpha}{Dis(w_1, w_2) + \alpha} \quad (1)$$

In (1), α is an adjustable parameter, in this system, we set α as 1. Considering the forest structure of this synonyms library, we calculate the distance by shortest path between two word sememe's nodes. The similarity could be then generated by the formula above.

4 Experimental Results and Analysis

In this paper, we use web crawler to attain massive video descriptions as test data. These data come from 4 main Chinese video service websites and video themes involve in sports, news, movies and variety shows. By counting recall rate and precision rate, we could clearly illustrate the comparison of three level's structured strings' recall rate and precision rate in Table 4 and Table 5. Results of None Expanded Mode and Expanded Mode are also exhibited below.

Table 4. Recall rate

Strings' Levels	None Expanded	Both Expanded
0-Level	0.6094	0.6322
1-Level	0.6911	0.7311
3-Level	0.7455	0.7844

Table 5. Precision rate

Strings' Levels	None Expanded	Both Expanded
0-Level	0.6250	0.6417
1-Level	0.6350	0.6583
3-Level	0.6756	0.8316

From the statistics above, we can see that the higher the string's level is, the higher the recall rate and precision will be. In addition, synonyms and hypernyms expanded mode could provide a better performance than none expanded mode do. It is natural that users' experience will be enhanced by structured semantic strings' expression and its expansion results.

5 Conclusion and Future Works

We propose a new method to semantically retrieve video description in this paper. On one hand, after NLP parsing each sentence in documents, the structured strings containing grammar information will be indexed in the inverted indexing files. On the other hand, searching contents would experience NLP parsing and then be disassembled and expanded, when searching contents aggregate again as a searching set, it could be submitted and help complete a better searching effect. The approach has been applied in our video semantic retrieval system Xunet and performs well. Also,

the experiment tested on Xunet demonstrates the advance of our method by concrete statistics.

Admittedly, there exist some limitations of our approach. For example, sentences which can't be processed by NLP will be discarded and it may cause losing information, in addition, too large quantity of expansion will influence the searching effect. We will concentrate on these problems to balance the semantic expressing and searching efficiency in the future.

Acknowledgments. Our work was supported in part by the National Key Technology R&D Program (2008BAH28B04), "A New Generation of Video and Television Service System with Support for Cross-regional and Multi-operator". What's more, project "Content Distributed Service Commercialization Based on IPV6" (CNGI-09-03-14) hold by National Development and Reform Committee sponsored our work in the later period.

References

1. Lee, D.L., Hueli, C., Seamons, K.: Document Ranking and the Vector-space Model. *IEEE Transaction on Software* 14, 67–75 (1997)
2. Chang, S.-F., Chen, W., Meng, H.J., Sundaram, H., And, D.: A Fully Automated Content-Based Video Search Engine Supporting Spatiotemporal Queries. In: *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 602–615 (1998)
3. Baştan, M., Çam, H., Güdükbay, U., Ulusoy, O.: BilVideo-7: An MPEG-7 Compatible Video Indexing and Retrieval System. *IEEE Transaction on Multimedia* 17, 62–73 (2010)
4. Campbell, M., Haubold, A.: IBM research TRECVID-2006 video retrieval system. In: *TREC Video Retrieval Evaluation Proceedings* (2006)
5. Duan, L.-Y., Xu, M., Tian, Q., Xu, C.-S.: A Unified Framework for Semantic Shot Classification in Sports Video. *IEEE Transactions on Multimedia* 7(6), 1066–1083 (2005)
6. Shady, S., Fakhri, K., Mohamed, K.: Enhancing Search Engine Quality Using Concept-based Text Retrieval. In: *International Conference on Web Intelligence, IEEE/WIC/ACM*, pp. 26–32 (2007)
7. Ballan, L., Bertini, M., Del Bimbo, A., Serra, G.: Video Annotation and Retrieval Using Ontologies and Rule Learning. *IEEE Transaction on Multimedia* 17, 80–88 (2010)
8. Lux, M.: Caliph & Emir: MPEG-7 photo annotation and retrieval. In: *Proceedings of the Seventeen ACM International Conference on Multimedia*, Beijing, China, pp. 925–926 (2009), <http://www.semanticmetadata.net/features/>
9. Harbin Institute of Technology LTP website, <http://ir.hit.edu.cn/>
10. Zheng, Q., Zhou, Z.: An MPEG-7 Compatible Video Retrieval System with Support for Semantic Queries. In: *International Conference on Consumer Electronics, Communications and Networks (CECNet)*, pp. 1035–1041 (2011)
11. Xu, S., Zhu, L.-J., Qiao, X.-D., Xue, C.-X.: A Novel Approach for Measuring Chinese Terms Semantic Similarity based on Pairwise Sequence Alignment. In: *Fifth International Conference on Semantics, Knowledge and Grid*, pp. 92–98 (2009)

Color Image Segmentation Algorithm Based on Affinity Propagation Clustering

Lei Wang and Lin Zhang

Institute of Modern Education Technology, Shangluo Univeristy,
Shangluo,726000, China
Sfch007@gmail.com, zlin002@163.com

Abstract. This paper proposes a color image segmentation algorithm based on affinity propagation clustering. Firstly, the color image is converted to other color space; and then color image is sampled, and the sampling data is clustered by affinity propagation algorithm with given the number, then we can get their class label of the rest data according to the maximum similarity rules; finally, using the regional combined methods ,we can remove independent small regions, and get revised image segmentation result. The simulation experiment shows that this method can achieve the application of affinity propagation algorithm in large-scale color image segmentation. And it has a faster processing speed and satisfactory segmentation results, more in line with the human whole visual perception.

Keywords: Affinity Propagation, Data Sampling, Image Segmentation, Regional Merger.

1 Introduction

Image segmentation is the basis for computer vision processing technology, and also the basic problem of pattern recognition and image processing. So far there is not a general segmentation method, nor is there an objective rule to judge the success of segmentation. Therefore, the research of new image segmentation technique becomes very practical.

Frey & Dueck in the journal Science in 2007 proposed a new clustering algorithm: Affinity Propagation (AP) [1]. Its main advantage reflects in the fast speed when handling datas which have a lot of class. AP algorithm has been proved more effective than K-means algorithm[1]. For example, AP algorithm only cost 5 min can find accurately a small amount of pictures ,which can explain all kinds of handwriting type from thousands of handwritten postal code pictures. But K-means algorithm takes 500 million years to achieve the same precision[2]. The algorithm has been used in the following fields: searching face image, finding gene exon, searching optimal route and so on. But the traditional AP algorithm has still some problems in practical applications, such as: Since the algorithm's time complexity and space complexity is a serious subject to number of sample data N , it causes the traditional AP algorithm not to handle large-scale data, especially image segmentation.

In this paper, we propose a color image segmentation algorithm based on affinity propagation clustering. It can be divided into three parts: (1) Pretreatment process for color image. In this section , RGB color image is converted to L*u*v* color image, whose aim is to remove the correlation of the three RGB components, and use the L*u*v* color space which conforms more to the human visual perception; (2) Image segmentation based on affinity propagation (AP) clustering. It includes other three parts: data sampling, APGNC clustering and obtaining class label of other data points. It is the aim to obtain the whole image segmentation results; (3) Posttreatment process for segmentation results. It can modify segmentation results using regional merging methods, and the aim is to get rid of wrong points or miss points (independent region) in target region, and ensure the integrity of the target.

The rest of this paper is organized as follows: Part II introduces the pre-treatment process for color image; Part III introduces image segmentation based on affinity propagation(AP) clustering; Part IV introduces the post-treatment process for segmentation result; Part V shows experimental results and analysis; Finally in Part VI we will give the conclusions.

2 Pretreatment Process for Color Image

Compared with gray images, color component is the unique feature of color images. If the brightness and color components of color image can be described separately, color images will be more accurately expressed. Color space conversion is one of primary means to achieve these goals.

Depending on the different application purposes, color space can have many different forms, such as RGB, HSL, HSV, CIEL*u*v*, CIEL*a*b*, and so on. Currently the most widely used color space is RGB space. Although it is simple and don't need conversion, it has a high correlation between the three components R, G, B. It is detrimental to image segmentation. For color images, the color space conversion is to convert RGB color image to a particular color system, and then we can analyse and process color image.

CIEL*u*v* color space is constructed according to human visual system. L* represents brightness; The u* and v* jointly represent a certain color, including some color which physical world can't recreate. In addition, the tonal line and saturation line of CIEL*u*v* color space are most similar to linear transformation, and it can well improve the visual non-formal problems. Therefore, using CIEL*u*v* color space to express brightness and color information is very ideal. CIEL*u*v* space is based on the XYZ space via a nonlinear transform [3]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

CIEL*u*v* is calculated as follows:

$$\begin{aligned} L^* &= 116(Y/Y_0)^{1/3} - 16 \\ u^* &= 13L^*(u' - u_0) \\ v^* &= 13L^*(v' - v_0) \end{aligned} \tag{2}$$

Among them,

$$\begin{aligned}u' &= \frac{4X}{X+15Y+3Z} \\v' &= \frac{6Y}{X+15Y+3Z}\end{aligned}\tag{3}$$

X_0, Y_0, Z_0 are the corresponding three color components of white in XYZ color space; Y_0, u_0, v_0 are the corresponding three color components of white in CIEL*u*v* color space.

3 Image Segmentation Based on AP

To reduce the space complexity and time complexity of affinity propagation, there are usually two ways: (1) the image is divided into some non-overlapping subgraphs, and then these subgraph as a data point are clustered separately [4]; (2) using sparse affinity propagation clustering[5]. Although both methods are able to reduce a certain space complexity and time complexity of affinity propagation, the algorithm itself complexity is higher, and not easy to come true [6].

Under a regular PC (Matlab environment), the number of data to be clustered by AP algorithm is generally limited to less than 3000 [7], and the corresponding image is also limited to 60*60 or less. But using the method in this paper, with selecting the appropriate sampling rate, we can greatly expand the application of AP algorithm in image segmentation.

For color images via pretreatment processing (color space conversion), we execute the initial AP clustering in this section. This part can be divided into three steps: First, we sample the input color image (including automatic uniform sampling and manual designated sampling). The aim is to reduce the number of data processing, and speed up the clustering; then, we execute affinity propagation with given number of clusters (APGNC) for sampling data, to obtain the corresponding cluster center. The aim is to be ready to calculate class label of other data points in the next step; finally, according to the maximum similarity rules, we can calculate the similarity between other data points which are not sampled and cluster centers in sequence, and get their class labels. The aim is to extend cluster results to the whole image.

3.1 Data Sampling

- Automatic uniform sampling. Set the sampling rate "rate", its scope:[0,1]. Sampling the whole image by uniform methods, which would be sampled every interval "1/rate" points. Obviously, the higher the sampling rate, the less sampled data, so the shorter the running time of clustering, the worse clustering results; and vice versa. Therefore, considering the efficiency and effectiveness of the algorithm, the number of sampled data can be limit between 300 to 2000 by changing the sampling rate, in order to ensure that the clustering process can be completed within an acceptable time.

- Manual designated sampling. Because the automatic uniform sampling is based on sampling rate, when sampling rate is smaller, the interval (1/rate) between the sampling data is larger. So we may ignore or skip directly some small target, which is obviously not conducive to get the better sampling data. Therefore, when the sampling rate is less than a value (default 1%), it should be combined with manual designated sampling, in order to add some interest small target region or point.

3.2 Affinity Propagation with Given Number of Clustering (APGNC)

We must firstly calculate the similarity S and performance P between these sampling data points. Then, according to user-defined number of final clusters (DK), we execute AP algorithm with given number of cluster (APGNC), and get a group of cluster centers. After this processing, we get the satisfactory clustering results.

As we know, AP algorithm cannot directly predict the final number of clusters before clustering. In other word, the traditional AP can't get the clusters results with given number of clustering like K-means. But according to the idea of the literature [8], we can achieve these goals by using the dichotomy judgment.

Here we give a detailed description of APGNC:

Algorithm: APGNC (Affinity Propagation with Given Number of Cluster)

Steps:

- ① Initialize AP: set the final number of clusters to DK , the performance $P \leftarrow \text{median}(S)$, the upper bound $P_high = 0$, the lower bound $P_low = 2 * P$.
- ② Put the performance P into AP algorithm, then run AP algorithm once: get the number of clusters K .
- ③ Judge as follows: If $K > DK$, it shows the number of clusters is too many, so we need to reduce the P -value: $P_high \leftarrow P$ and $P_low \leftarrow 2 * P_low$, get a new performance $P \leftarrow 1/2 * (P_low + P_high)$, turn into ②; if $K < DK$, it shows the number of clusters is too little, so we need to increase the P value: $P_low \leftarrow P$, get a new performance $P \leftarrow 1/2 * (P_low + P_high)$, turn to ②; if $K = DK$, it shows that we achieve the goal, turn into ④.
- ④ Output corresponding clustering results, and the algorithm stops.

3.3 Obtaining Class Label of Other Data Points

After APGNC clustering, sampling data points have already got their own class label, and the corresponding cluster centers, while the rest of other data has no their class labels. In this section we will expand the initial clustering results to the whole image as follows:

We can calculate the similarity between nonsampled data points and cluster centers in sequence, according to the maximum similarity rules, then respectively add their class label C_x :

$$C_x = \arg \max \{-\|x - C_i\|\} \quad (4)$$

C_i is the i -th cluster center after AP clustering, $i = 1, 2, \dots, DK$.

4 Posttreatment Process for Segmentation Results

Through the above steps, affinity propagation clustering has already completed color image segmentation. But because AP is sensitive to data points, sometimes there has a large number of small independent region, when the color image is clustered. This seriously affected the whole effect of segmentation, and reduced segmentation accuracy, and doesn't conform to human whole visual perception. Therefore, we need to merge these independent small regions by using region merging algorithm of the morphology, thus amended the final segmentation results.

Because of AP algorithm only using color information, not pixel spatial information, it is easy to produce some smaller independent regions. Region merging is to merge two adjacent region in space, which are most similar in color. The aim is to remove trivial region and noise point, and ensure the integrity of the target. In this paper we mainly use the acreage and color features of regions to region merging [9]. Then, we will give our region merging rules in this paper as follow:

Firstly, establish region adjacency table (RAT) for the whole region, and set the acreage of the region to threshold T. Define the region as the object region i , whose acreage is less than the threshold T.

Secondly, find out all adjacent regions of the object region i by scanning region adjacent table (RAT). Then calculate the similarity between the object region i and all adjacent regions. And then according to the maximal similarity criteria (color distance), find out the most similar (minimum distance) region as the target region j , which will be merged with the object region i . Here we use the color distance as the similarity of the region as follows:

$$D_{ij}^c = \frac{|r_i| \cdot |r_j|}{|r_i| + |r_j|} \cdot \|\bar{\mu}_i - \bar{\mu}_j\|^2 \tag{5}$$

$|r_i|, |r_j|$ represent the number of pixels (acreage) in the i -th and j -th region; $\bar{\mu}_i, \bar{\mu}_j$ represent the mean color of two regions; $\|\bullet\|$ represents Euclidean distance.

Finally, change the class label of the object region i into the class label of the target region j , and update the region adjacency table (RAT). Do like this until all regions scanned.

The following is the region merging rules in this paper:

1) Initialization. n is the initial number of region. $R = \{r_1, r_2, \dots, r_n\}$, and $R' = \Phi$. R is the set of regions before region merging. R' is the set of regions after region merging. r_m is the m -th region, $m = 1, 2, \dots, n$.

2) If the acreage of $r_i <$ the threshold T, then r_i is the object region, $R = R - r_i, i = 1, 2, \dots, n$.

3) $r_j = \arg \min(Dis(r_i, r_j) | 1 \leq j \leq n, r_j \in R)$, then r_j is the target region.

- 4) The Class label of $r_i \leftarrow$ the Class label of r_j , and $r' = r_i \cup r_j$,
 $R = R - r_j, R' = R' + r'$
- 5) Repeat 2) ~ 4), until R is empty

5 Experimental Results and Analysis

5.1 Experimental Data

To verify the accuracy and practicality of the proposed method in this paper, we select three test images (Figure 1) from the Berkeley color image database. This is because each image in Berkeley color image database has a corresponding manual segmentation results, and it is easy to compare our experimental results to manual segmentation results. The detailed image information can be gotten in the Table 1.

For each test images, the similarity of APGNC algorithm is set to the negative Euclidean distance, and the damping coefficient “damp” is set to 0.7, and the initial performance P is set to median (S), and the given number of cluster “DK” is set to the number of segmentation like Table 1, and sampling rate is set to 2%, and the acreage threshold T of region merging is set to 200. All experiments are executed in the same environment (CPU: Intel Core (TM) 2,1.86 GHz With 1G memory), experimental software: MATLAB 7.8.0 (R2009).



Fig. 1. Experimental Image (bird, airplane, church)

Table 1. Image Information

<i>Experimental Image</i>	<i>Size</i>	<i>Color Space</i>	<i>Number of segmentation</i>	<i>From</i>
Birds	321*481	RGB	2	Berkeley Database
Airplane	321*481	RGB	2	Berkeley Database
Church	321*481	RGB	3	Berkeley Database

5.2 Experimental Results

In order to verify the effectiveness and practicality of this method, we also joined the RGB color image segmentation based on K-means cluster, and its number of clusters is set to the specified number of segmentation like Table 1, and then we randomly choose color pixel as the initial clustering center.

In Figure 2, we find that the APGNC algorithm based on data sampling can complete color image segmentation in most cases, which can't be accomplished by traditional AP algorithm. In addition, we can change the value of sampling rate, so that the algorithm is used for larger color image segmentation. Obviously, the smaller the sampling rate, the less sampling data, and the faster APGNC cluster, more easy to overlook the useful information, the worse segmentation results; and vice versa. Also, color image segmentation results based on the K-means have a lot of small independent regions, which seriously affect the accuracy of the segmentation. And compared with the K-means, the proposed color image segmentation method in this paper can effectively reduce small isolated regions of segmentation results, in order to improve the clustering accuracy, and make the results to be more in line with the human whole visual perception.

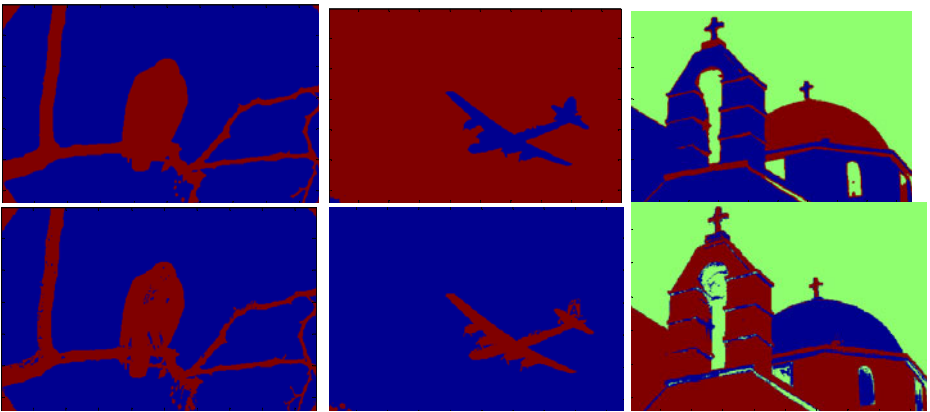


Fig. 2. Results of Clustering

(The first line is this method of clustering results;the second line is Kmeans clustering of clustering results)

In Figure 3, compared with manual segmentation results in the Berkeley image database, it is not difficult to find that the segmentation result based on our method is more in line with the classification results of the human eye, which is consistent with the facts; especially because of using a region merging strategy, we can fix the segmentation results of APGNC clustering to a certain extent, so that we remove small independent regions in the background region. But this drawback is obvious, because it can increase the time complexity of image segmentation, and image segmentation results becomes worse for complex background region, especially texture abundant image.



Fig. 3. Results of Segmentation

(The first line is this method of segmentation results, the second line is Kmeans clustering of segmentation results, the third line is manual segmentation results from Berkeley)

In Table 2 we separately lists the running time(s) of the proposed method and K-means. Obviously, this method has a faster running time, and can achieve more satisfactory segmentation results.

Table 2. Comparison Test

<i>Experimental Image</i>	<i>Running Time(s)</i>	
	<i>This Method</i>	<i>K-means</i>
Birds	9.2	36.6
Airplane	11.3	18.1
Church	11.1	11.6

6 Conclusion

In this paper, for the existing problems of affinity propagation clustering algorithm in large-scale data processing, we proposed color image segmentation method based on affinity propagation clustering by using color space conversion ,data sampling, and region merging, to further improve the image segmentation results. The simulation shows that image segmentation results using this method are more consistent with

human whole visual perception. However, this method still exist obvious deficiency that it does not use the texture and other useful spatial information of image but only color information. Therefore, how to use these useful spatial information and apply them to the clustering algorithm, is the research direction we should be continue in the next step.

Acknowledgments. The research work described in this paper was supported by the grants from the Shaanxi Province Education Science Planning Project of China (Project No.SGH0902197). And we would like to thank Prof. Wang Xili of Shaanxi Normal University for suggestions of this paper.

References

- [1] Frey, B.J., Dueck, D.: Clustering by Passing Messages Between Data Points [J]. *Sinence* 315(5814), 972–976 (2007)
- [2] Kelly K.: Affinity propagation slashes computing times [OL] (October 25, 2007), <http://www.news.utoronto.ca/bin6/070215-2952.asp>
- [3] Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J.: Color image segmentation: advances and prospects. *Pattern Recognition* 34(12), P225–P2281 (2001)
- [4] Frey, B.J., Dueck, D.: Mixture modeling by affinity propagation. *Neural Information Processing Systems* 18, 379–386 (2006)
- [5] Xiao, J.X., Wang, J.D., Tan, P., et al.: Joint affinity propagation for multiple view segmentation. In: *Proceedings of 2007 IEEE 11th International Conference on Computer Vision*, Institute of Electrical and Electronics Engineers Inc., Piscataway (2007)
- [6] Zhangren, Y., Zhao, H., Lu, X., Cao, M.: Spread of clustering based on similarity of the gray image segmentation. *Naval University of Engineering*
- [7] Gu, R.J., Wang, J.-C., Chen, G., Chen, S.: Neighbors for the dissemination of large data sets clustering. *Computer Engineering* (December 2010)
- [8] Wang, L., Wang, X.-L., Liu, G., Zhao, L.: A combination of semi-supervised pro and dissemination of improved adaptive clustering. *Computer Science* (December 2010)
- [9] Cheng, H.D., Jiang, X.H., Wang, J.: Color image segmentation based on homogram thresholding and region merging. *Pattern Recognition* 35(2), 373–393 (2002)

Part III

**Cognitive Science and
Brain-Computer Interface**

A EEG-Based Brain Computer Interface System towards Applicable Vigilance Monitoring

Hongfei Ji, Jie Li, Lei Cao, and Daming Wang

Department of Computer Science & Technology, Tong Ji University,
Cao an Highway No. 4800, Shanghai, China
jhff@tongji.edu.cn, nijanice@gmail.com, scallray@163.com,
damingwang.tongji@gmail.com

Abstract. Monitoring the changes of vigilance is very important, because decline in the operator's alert is one of the primary causal factors for many accidents. Electroencephalogram (EEG) based brain computer interface (BCI) for vigilance analysis and estimate have been rapidly advanced in recent years. However, there still exist many difficulties and challenges in their application. In this paper, we propose an applicable EEG-based BCI system for vigilance monitoring: the number of record channels could be reduced greatly with considering individual variability in EEG, since the practical BCI is expected to be implemented with a small number of channels. Furthermore, the vigilance model can be constructed according to various application situation. With the dry electrodes we are now developing, the system will be conveniently applied to vigilance monitoring in the future.

Keywords: Electroencephalogram, Brain Computer Interface, Vigilance Monitoring.

1 Introduction

Monitoring vigilance is a hot field since it is extremely useful to prevent some accident in performing attention-demanding and monotonous tasks[1-5], such as vehicle driving. Electroencephalographic (EEG) is assumed to be one of the most reliable indicators of vigilance and reflect human vigilance state, because it is the recording of electrical activity produced by the firing of neurons within the brain, and contains the objective brain state information[6]. Especially, the changes in EEG spectrum related to the vigilance state closely[7-10].

An EEG based brain computer interface (BCI) provides a direct way for the communication between the human brain and the computer [11]. There are some studies regarding BCI system for vigilance analysis[12-13], however, there still exist many difficulties and challenges in their application. Most of the earlier studies using EEG record in the a lots of channels to assessment of changes in cognitive states, while the practical BCI is expected to be implemented with a small number of channels. The relatively large individual variability in EEG and different application also result in the difficulty in vigilance analysis and estimate.

In this paper, we propose an applicable EEG-based BCI system for vigilance monitoring: the number of channels could be reduced greatly with considering individual

variability in EEG. Furthermore, the vigilance model can be constructed according to various application situation. With the dry electrodes we are now developing, the system will be conveniently applied to vigilance monitoring in the future.

The rest of the paper is organized as follows. The proposed BCI scheme in our study is introduced in Section 2, Section 3 describes the method, and section 4 presents corresponding result. Section 5 offers the conclusion.

2 System Scheme

The goal of the proposed system is to provide a practical BCI easy to use, the block diagram of the proposed BCI system is illustrated in Fig.1: Firstly, multi-channels raw EEG signals are collected in our experiment, and then some acquired EEG data in two definite vigilance states are selected. Then the spectral power in the four bands, δ (0-4 Hz), θ (4-8 Hz), α (8-13 Hz) and β (13-35Hz), and their ratios which are $(\theta + \alpha)/\beta$, α/β , $(\theta + \alpha)/(\alpha + \beta)$, and θ/β , are calculated as features for vigilance detection. After obtaining features in the previous step, Fisher score is utilized to further remove redundant features, and reduce indiscriminative channels. Finally, the vigilance model could be constructed by GMM method. Thereafter, the corresponding vigilance levels could be obtained.

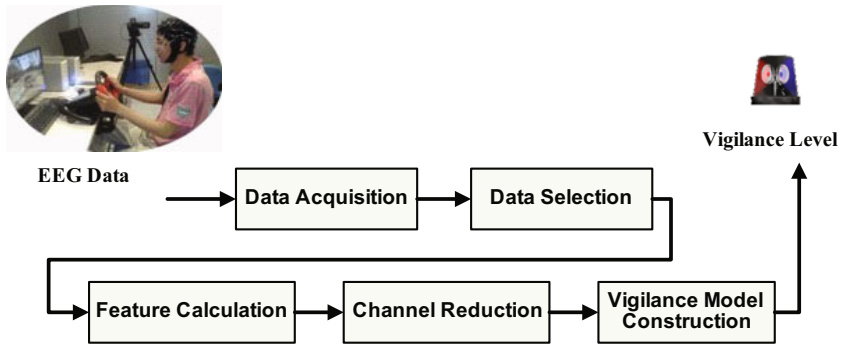


Fig. 1. Block diagram of the proposed BCI system

3 Methods

3.1 Data Acquisition

Multichannel EEG data were acquired by an ESI-128 channel high-Resolution EEG/EP systems (SynAmps2, Neuroscan) in the Lab for Brain Cognition and Intelligent Computing, Tong Ji University, China. The electrodes were placed based on the standard international 10-20 system. Nine healthy male students took part in the experiment. Since each experiment was taken after lunch, most of subjects would gradually get drowsiness even fall into a sleep rapidly, and therefore EEG data in different vigilance states were obtained. In order to distinct the subjects vigilance states especially in sleepy

and wakeful states, two channels' Electrooculogram (EOG), experiment video and the subject feedback information were all recorded simultaneously. The sample frequency was 1000Hz, and then down-sampled to 250Hz and bandpass filtered between 0.1 Hz and 40 Hz, and 5s time window was applied in the following calculation.

3.2 Channel Reduction

The channel reduction interface in the proposed BCI system is showed in Fig.2. In this interface, according to the experiment video, EOG and the subject feedback, we can select two contrastive data segments which represent two different ultimate vigilance states, that is, wakefulness and drowsiness for channel reduction analysis. Due to the close relationship between the EEG spectrum and the subject's vigilance state, the rhythm activities, that is, EEG power in the four specified bands, δ (0-4 Hz), θ (4-8 Hz), α (8-13 Hz) and β (13-30Hz) and their ratios, which are θ/β , α/β , $\theta + \alpha/\alpha + \beta$, and θ/β , are calculated as features. And for those features, Fisher score for each channel would be calculated, and utilized to further remove redundant features, and also reduce indiscriminative channels. The selected channels would be listed and illustrated with discriminative weights in the right panel to guide the channel selection in practical application.

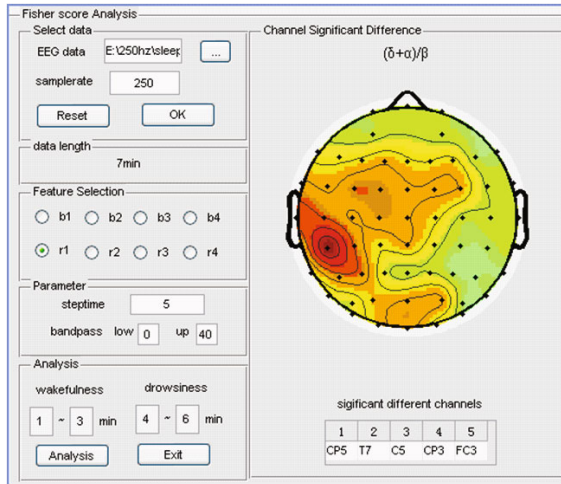


Fig. 2. Channel reduction interface

For each channel, its fisher score is computed to find the significant difference between wakefulness and drowsiness, which is given by

$$Fr = \frac{S_B}{S_W} \tag{1}$$

where S_B is the between-class covariance and is given by

$$S_B = (m_2 - m_1)^2 \quad (2)$$

where $m_1 = \frac{1}{N_1} \sum_{n \in C_1} x_n$, $m_2 = \frac{1}{N_2} \sum_{n \in C_2} x_n$. S_W is the between-class covariance matrix and is given by

$$S_W = \sum_{n \in C_1} (x_n - m_1)^2 + \sum_{n \in C_2} (x_n - m_2)^2 \quad (3)$$

C_1 represents the data at the wakeful state and C_2 represents the data at the drowsiness state. N_1 is the total number of the data points at the wakeful state and N_2 is the total number of the data points at the drowsiness state. x_n is the data point and $n = 1, 2, \dots$

For each channel, its fisher score is computed, and then channels with the some largest Fisher scores are retained as the most significant channels. By this way, the most significant channels for vigilance discrimination are selected, and the number of channels in application can be greatly reduced.

3.3 Vigilance Model Construction

Based on the selected channels' features, a popular clustering method, gaussian mixture model-based (GMM) clustering, is applied to assess a gradual vigilance changing process, and obtain a vigilance model.

The algorithm of Gaussian mixture model clustering is using Gaussian mixture model (GMM) for data clustering. The GMM is a weighted sum of M component Gaussian densities represented by the following equations:

$$p(x/\lambda) = \sum_{i=1}^M \omega_i g(x/\mu_i, \Sigma_i); \quad (4)$$

where,

$$g(x/\mu_i, \Sigma_i) = \frac{1}{((2\pi)^{D/2} |\Sigma_i|^{1/2})} \exp - \frac{1}{2} (x - \mu_i)' \Sigma_i^{-1} (x - \mu_i); \quad (5)$$

here x is the sample data with D -dimensional features, and $g(x/\mu_i, \Sigma_i)$ is a Gaussian density function whose mixture weight is ω_i , while $\lambda = \{\omega_i, \mu_i, \Sigma_i\}$ is a collective parameter with μ_i and Σ_i denoting the mean vector and the covariance matrix respectively. Given the GMM configuration, the next vitally important step is to estimate the parameters of the GMM, λ , which can be realized through maximum likelihood (ML) estimation. However, ML parameter estimates can be obtained iteratively using a special case of the expectation maximization (EM) algorithm. The GMM is often used for data clustering due to its ability to form smooth approximations for arbitrarily shaped densities so as to represent various sample distributions, and therefore, we use it to assess a gradual vigilance changing process.

The implementation of vigilance model analysis is comprised of two parts: 1) the parameter estimating with iterative EM approach based on GMM and 2) the classifying of testing data. During the period of parameter estimating, the GMM configuration is constructed and the initial cluster number is set to M firstly. Then the parameters of the GMM which include the collections of λ are iteratively estimated through the

algorithm of expectation maximization (EM) until the cluster number of M reduces to K that is desired. This moment, parameters of K clusters including the mean vector and covariance are obtained respectively. When it comes to the episode of testing, the collections of derived from the part of parameter estimating can be utilized to determine whether the testing data belongs to one type of cluster or not. According to various application situation, we can assign different values to the desired cluster number K . For example, for the situation required a sophisticated classification of vigilance, the K can be set to a large value. On the contrary, we can set K to '2', if only a rough vigilance classification is required.

4 Results

Fisher score analysis is applied to find the significant channels between wakeful state and drowsy state in above 8 features. Table 1 shows the classification result of 9 subjects by using the 5 channels and 62 channels data respectively. The result indicates, for most subjects, although the accuracy by using the 62 channels' data are higher, the 5 channel could also acquire high accuracy, especially in the feature α/β , the larger fisher score may represent the higher classifying accuracy in the corresponding channel for most subjects. Fig.3 and Fig. 4 illustrate 3-class cluster results for about thirteen

Table 1. The SVM classification accuracy of 9 subjects by using the 5 channels' and 62 channels' data respectively. For each subject, the left column denote 5 channels result, and the right denotes 62 channels.

Sub(%)	1	2	3	4	5	6	7	8	9									
δ	54.5	81.8	85.5	100	89.9	98.6	64.7	96.1	96.6	100	86.2	100	94.4	100	75.0	90.0	83.3	100
θ	48.5	53.0	96.4	100	87.0	100	80.4	100	79.3	100	89.7	100	100	100	80.0	95.0	100	100
α	72.7	97.0	100	100	89.9	100	84.3	98.0	100	100	93.1	100	100	100	85.0	90.0	100	100
β	54.5	100	100	100	97.1	100	90.2	100	100	100	93.1	100	100	100	70.0	100	100	100
$r1$	59.1	92.4	92.7	100	79.7	100	80.4	100	100	100	96.5	100	94.4	75.0	100	100	100	100
$r2$	72.7	89.3	85.5	100	95.7	100	68.6	98.0	100	100	100	100	100	100	85.0	100	100	94.4
$r3$	56.1	71.2	61.8	100	88.4	87.0	90.2	98.0	100	100	79.3	100	55.6	100	64.0	100	77.8	94.4
$r4$	60.6	68.1	72.7	100	82.6	94.2	92.2	100	100	100	79.3	96.6	88.9	88.9	85.0	100	83.3	83.3

minutes 5 channels' and 62 channels' EEG data respectively. EEG data acquired during a wakefulness(1-40 time windows)-sleepiness(81-120)-wakefulness(121-160) period. Based on the vigilance model calculated in training stage, the vigilance levels in testing stage are ranked reasonably. Experiment results show that, compared with using the 62 channels' data, the 5 channels' model identifies more reasonable vigilance states. It can be concluded that the channel reduction probably removes some redundant features, and avoids over fitting.

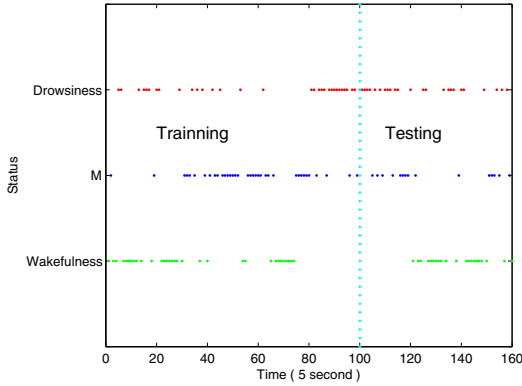


Fig. 3. Vigilance estimate results for 62 channels EEG data acquired during a wakefulness-sleepiness- wakefulness period. The broken line denotes the partition between the training data and testing data.

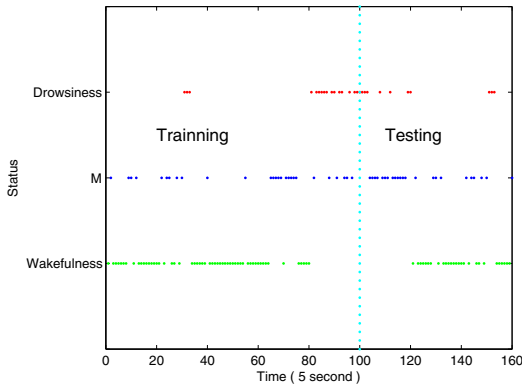


Fig. 4. Vigilance estimate results for 5 channels EEG data acquired during a wakefulness-sleepiness- wakefulness period. The broken line denotes the partition between the training data and testing data.

5 Conclusion

Designing BCI system to maximize their usability and portability in practical operational environments has attracted increasing attention. With subject-adaptive channel reduction and situation-adaptive vigilance model construction, the proposed system can enhance the usability of EEG based vigilance monitoring system. Now, we are also developing a wireless embedded BCI system, which consists of a dry electrode EEG data acquisition and amplification unit, a wireless transmission unit, a signal-processing unit. The method in the proposed system can also be employed to enhance the efficiency of the embedded system.

Acknowledgment. The work was supported by the Science and Technology Commission of Shanghai Municipality, China (Grant No. 09511502500), and the National Natural Science Foundation of China (Grant No.61105122).

References

1. Wright Jr., K.P., Hull, J.T., Czeisler, C.A.: Relationship Between Alertness, Performance, and Body Temperature in Humans. *American Journal of Physiology* 283, 1370–1377 (2002)
2. Oken, B.S., Salinsky, M.C., Elsas, S.M.: Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology* 117 (2006)
3. Vural, E., Cetin, M., Ercil, A., Littlewort, G., Bartlett, M., Movellan, J.: Drowsy driver detection through facial movement analysis. *Human-Computer Interaction* 247, 6–18 (2007)
4. Liang, S.F., Lin, C.T., Wu, R.C., Chen, Y.C., Huang, T.Y., Jung, T.P.: Monitoring Drivers Alertness Based on the Driving Performance Estimation and the EEG Power Spectrum Analysis. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 6, pp. 5738–5741 (2005)
5. Balasubramanian, V., Adalarasu, K., Gupta, A.: EEG based analysis of cognitive fatigue during simulated driving. *International Journal of Industrial and Systems Engineering* 7, 135–149 (2011)
6. Artaud, P., Planque, S., Lavergne, C., Cara, H., de Lepine, P., Tariere, C., Gueguen, B.: An on-board system for detecting lapses of alertness in car driving. In: *Proceedings of the Fourteenth International Conference on Enhanced Safety of Vehicles*, Munich, Germany (1994)
7. Lin, C.T., Wu, R.C., Jung, T.P., Liang, S.F., Huang, T.Y.: Estimating Driving Performance Based on eeg spectrum analysis. *EURASIP Journal on Applied Signal Processing* 19, 3165–3174 (2005)
8. Cao, L., Li, J., Sun, Y.R., Zhu, H.P., Yan, C.G.: EEG based Vigilance Analysis by Using Fisher Score and PCA Algorithm. In: *The IEEE proceedings of the International Conference on Progress in Informatics and Computing Conference* (2010)
9. Li, J., Yan, C.G., Xia, B., Cao, L., Wang, D.M., Sun, Y.R.: EEG-based Brain Computer Interface for Vigilance Analysis and Estimate. In: *Conference on Asia-Pacific Signal and Information Processing Association, APSIPA ASC* (2010)
10. Shi, L.C., Yu, H., Lu, B.L.: Semi-supervised clustering for vigilance analysis based on EEG. In: *International Joint Conference on Neural Networks*, pp. 1518–1523 (2007)
11. Vidal, J.J.: Toward direct brain Ccomputer communication. *Annu. Rev. Biophys. Bioeng.* 2, 157–180 (1973)
12. Hsieh, H.Y., Liang, S.F., Ko, L.W., Lin, M., Lin, C.T.: Development of a Real-Time Wireless Embedded Brain Signal Acquisition/Processing System and its Application. In: *IEEE International Conference on Systems, Man and Cybernetics, SMC 2006*, vol. 5, pp. 4374–4379 (2006)
13. Lin, C.T., Chen, Y.C., Huang, T.Y., Chiu, T.T., Ko, L.W., Liang, S.F., Hsieh, H.Y., Hsu, S.H., Duann, J.R.: Development of wireless brain computer interface with embedded multitask scheduling and its application on real-time driver's drowsiness detection and warning. *IEEE Transactions on Biomedical Engineering* 55, 1582–1591 (2008)

Author Index

- Bai, Ting 469
Bellemans, Tom 249
Benítez, Ángela López 603
- Cai, Xianfa 29
Cao, Lei 365, 743
Centeno, Franxavier 603
Chang, Faliang 697
Chen, Dongfang 591
Chen, Kai 475
Chen, Niannian 641
Chen, Peijian 317
Chen, Qihua 557
Chen, Shuwei 159
Chen, Yijin 171
Cheng, Shu 469
Cheng, Yong 563
Cidade, Geraldo 569
Cools, Mario 249
Costa, Lilian 569
Cui, Baojiang 259
- Deng, Jingyu 527
Deng, Xiangyu 85
Ding, Zhe 3
Domènech, Carles 603
Dong, Meng-juan 267
Dong, Pinliang 131
Drias, Habiba 379
Du, Shaokun 355
Dun, Pu 77
- Fan, Minhu 501
Fan, Yong 641
Feng, Lifu 99
- Feng, Yipan 557
Franco, Camilo 307
- Gao, Lin 641
Gao, Meifeng 457
Gao, Ting 625
Gao, Yi 281
Gómez, Daniel 635
Guan, Sheng-Uei 275
Gui, Bing 55
Guo, Ling 355
Guo, Weizhao 13
- He, Jie 317
He, Yan 673
Herrera, Francisco 287
Hou, YingKun 563
Hu, Xiaopeng 715
Hu, Yun 151
Huang, Jie 483
Huang, Mengxing 507
Huang, Suling 575
- Janssens, Davy 249
Ji, Hongfei 743
Ji, Yang 427, 437
Jia, Bing 371
Jia, Yang 547
Jiang, Xiao-Long 219
Jin, Xin 527
- Katarzyniak, Radoslaw P. 105, 229
Kerre, Etienne 297
Kowalczyk, Ryszard 105, 229, 343

- Lei, Ying 391
 Li, Beibei 463
 Li, Chunlei 483
 Li, Cunhua 151
 Li, Fangmin 685
 Li, Jiangbao 697
 Li, Jie 29, 743
 Li, Qiyan 501
 Li, Wei 141
 Li, Xiaorong 641
 Li, Xuegang 507
 Li, Xuejin 239
 Li, Yufeng 641
 Li, Zhenghao 475
 Li, Ziqiang 267
 Li, Zuoyong 563
 Liang, Xiaobing 259
 Liao, Shizhong 191, 201
 Liu, Chengli 365
 Liu, Fei 275
 Liu, Feng 141
 Liu, Gang 99
 Liu, Guangde 335
 Liu, Jianguo 131
 Liu, Jingneng 445
 Liu, Jun 159
 Liu, Weiyi 63
 Liu, Ying 99
 Liu, Yong 191
 Liu, Zhi 557
 Liu, Zhining 489
 Liu, Zhoufeng 483
 Lorkiewicz, Wojciech 229
 Lu, Jie 297
 Lu, Yue 495
 Luo, Jiangfeng 93
 Luo, Pei 679
 Luo, Qijun 39
 Lv, Ruijuan 207

 Ma, Bo 613
 Ma, Shiwei 653
 Ma, Teng 679
 Ma, Yide 85
 Ma, Zhixin 3
 Macià, Joaquim Lloveras 603
 Mansour, Khalid 343
 Martínez, Luis 287
 Marturelli, Leandro 569
 Menacer, Djamel Eddine 379

 Meng, Cai 547
 Miao, Yubin 537
 Montero, Javier 307, 635
 Mu, Shaomin 327
 Mu, Xin 421, 705
 Muñoz, Susana 635

 Naderpour, Mohsen 297
 Nie, Shengdong 625
 Niu, Zhendong 239
 Nuo, Qun 77

 Pan, Xiang 557
 Pan, Zhen 475
 Pérez-Ràfols, Francesc 603
 Popek, Grzegorz 105

 Qi, Lu 69
 Qiao, Guohui 575
 Qin, Guofeng 501
 Qin, Tao 527
 Qionгда, Niluo 77
 Quan, Shuhai 469

 Rodríguez, J. Tinguaro 307
 Rodríguez, Rosa M. 287
 Ruan, Da 249

 Shen, Ruimin 281
 Shen, Yi 427, 437
 Shi, Bo 69
 Shi, Ying 469
 Sibertin-Blanc, Christophe 379
 Song, Shulin 705
 Sun, Bing 457
 Sun, Fengyan 613
 Sun, Lanlan 151
 Sun, Minglei 575, 581
 Sun, Qiaoyu 495
 Sun, Xiaoyan 697
 Sun, Xiwen 625
 Sun, Yi 715
 Sun, Zhijun 721
 Sun, Zhonghua 115

 Tang, Chuan 317
 Tang, Shichuan 575, 581
 Teng, Jinyu 23
 Tian, Bo 653
 Tian, Kailin 63

- Tian, Shengfeng 327
 Trashy, Nyima 77
 Tso, Yong 77
 Tu, Binbin 691

 Wang, Bing 335
 Wang, Daming 743
 Wang, Di 581
 Wang, Hailong 317
 Wang, Hong 69
 Wang, Jianxin 259
 Wang, Lei 731
 Wang, Qingfeng 121
 Wang, Sheng 685
 Wang, Song 721
 Wang, Ting 275
 Wang, Wei-Ye 219
 Wang, Xian 421, 705
 Wang, Yan 281
 Wang, Yuanjun 625
 Wang, Yuhui 121
 Wang, Zheng 99
 Wei, Jia 29
 Wen, Guihua 29
 Wen, Wen 581
 Wets, Geert 249
 Wu, Fanggu 575
 Wu, Jian-liang 335
 Wu, Wei 49
 Wu, Yiquan 427, 437
 Wu, Yuwei 613
 Wu, Zhuangzhi 679

 Xiao, Guoqiang 475
 Xu, Xiao 171
 Xu, Yang 159, 181, 207, 355
 Xu, Yusheng 3
 Xue, Xiaoliang 547

 Yan, Lei 489
 Yan, Wenjun 49
 Yang, Banghua 249
 Yang, Fan 281
 Yang, Lei 653
 Yang, Qiang 49
 Yang, Tingting 115
 Yang, Yan 715

 Yang, Yang 365
 Yang, Yongjian 371
 Yang, Yuexiang 317
 Yin, Chuanhuan 327
 Yin, Jian 13
 Yu, Fengqin 673, 691
 Yu, Wenjiao 507
 Yuan, Jing 721
 Yue, Kun 63

 Zeng, Guihua 445
 Zeng, Wei 365
 Zhang, Bin 581
 Zhang, Chaofei 715
 Zhang, Chunxia 239
 Zhang, Du 401
 Zhang, Fangsheng 421
 Zhang, Haiying 415
 Zhang, Hongxiang 39
 Zhang, Jun 371
 Zhang, Ke 527
 Zhang, Lin 731
 Zhang, Min 591
 Zhang, Rong 581
 Zhang, Shanhong 13
 Zhang, Weiming 93
 Zhang, Xiaofeng 495
 Zhang, Xiaoxin 23
 Zhang, Yan 421, 705
 Zhao, Ang 39
 Zhao, Li 181
 Zhao, Ning 201, 685
 Zhao, Zhi-hong 141, 463
 Zhao, Zhihui 201
 Zhaxi, Gama 77
 Zheng, Huicheng 665
 Zheng, Quan 721
 Zhong, Chong-quan 517
 Zhong, Xiaomei 159
 Zhou, Hongtao 365
 Zhou, Zhou 219
 Zhu, Cheng 93
 Zhu, Donghai 507
 Zhu, Qiang 537
 Zhu, Yan-dong 517
 Zhu, Yunlei 63
 Zong, Guanghua 575